

This is a postprint version of the following published document:

Migle Laukyte. 2019. AI as a Legal Person. In *Proceedings of International Conference on Artificial Intelligence and Law (ICAIL 2019)*. ACM, New York, NY, USA, 5 pages.

DOI: [10.1145/3322640.3326701](https://doi.org/10.1145/3322640.3326701)

© 2019 Association for Computing Machinery

AI as a Legal Person

Migle Laukyte
Private Law Department
Universidad Carlos III de Madrid
Getafe, Madrid, Spain
migle.laukyte@uc3m.com

ABSTRACT

The idea of the legal personhood of artificial intelligence (AI) – the idea that intelligent agents can have rights and incur obligations under the law – is controversial, and in fact is often dismissed out of hand: in this paper I will argue that, on the contrary, such legal personhood may be the next big challenge for our legal systems, and we need it to deal with the new kinds of complexity introduced by AI. Furthermore, I argue that we already have experiences we can look to: to this end we can draw on the reasoning applied to the legal personhood recognized for corporations and other nonhuman entities. In order to do this, I address some of the criticisms against ascribing legal personhood to AI. I also look at the Canadian and EU ethical guidelines so as to keep the development of AI within the framework of human values, and I show that an ascription of legal personhood to AI is consistent with them. I also address a few of the big issues involved in making the legal personhood of AI a reality.

CCS CONCEPTS

• Artificial Intelligence • law • governmental regulation

KEYWORDS

Artificial Intelligence, Legal Personhood, Corporation, Rights, Ethics

A

1 Introduction

The very idea of ascribing legal personhood to artificial intelligence (AI) usually attracts criticism. In this paper I will argue that, on the contrary, such legal personhood may well be the next big challenge for our legal systems, and that we already have models on which to build this new kind of legal personhood. I address the legal personhood of AI by arguing that it should be inspired by –but not copy-pasted from– the reasoning applied to the legal personhood recognized for corporations.¹

The discussion is laid out as follows. In Section 2, I briefly discuss what legal personhood means for corporations. In Section 3, I address the legal personhood ascribable to AI, considering a few arguments in favor and addressing the objections (3.1). In section 3.2, I explain why these debates should not hold back the idea of legal personhood for AI. In Section 4 I look at ethical considerations and ask whether legal personhood for AI is compatible with the Canadian Montréal Declaration for Responsible Development of AI (4.1) and the European Draft AI Ethics Guidelines for Trustworthy AI (4.2). Section 5 addresses what I consider to be the main problems in ascribing legal personhood to AI. Section 6 presents a few closing remarks.²

2 Legal Personhood

To readers without a legal background the concept of legal personhood might not be clear. We could see corporate legal personhood as a legal status accorded to a corporation, as a result of which the corporation gains certain rights (and obligations) that ensure its ability to operate as a going concern. The rights that corporations have vary a lot depending on the legal system, but they usually enjoy property rights, the right to sue and be sued, and the right to enter into contracts.

There are many theories on corporate legal personhood, such as the legal fiction theory, the concession theory, the group theory, the real entity theory, and the nexus of contracts theory [2], [3]. Among these, the one that admits of the idea of AI as a legal

¹ Corporations are not the only nonhuman legal persons, to be sure, but for the sake of clarity, and on account of limited space, I will be referring to corporations only.

² But, before we move on we need to fix a working definition of AI. In this paper AI is autonomous and intelligent enough to make its own decisions about its own actions and goals, and to figure out how to go about to achieve those goals without (or with very little and sporadic) human intervention. This kind of AI is also social

[1]: humans can interact with AI, and AI is able to read, interpret, and respond to human emotional behavior. This definition is possibly too broad, but what matters here to provide a broad picture of what it does.

person is the real entity theory, according to which corporation is more than just the sum of its individual human members' wills: it is an independent entity in its own right and has its own goals, and plans and continues its existence regardless of changes among its human members [2].³

3 Legal Personhood for AI

There is a question that needs to be asked: why should we accord legal personhood to AI in the first place? The answer is that we need to create a special kind of legal personhood for AI because developments in AI are showing that the current legal concepts of responsibility, product liability, and others are no longer fit for their overarching purpose, that of ensuring justice and protecting those whose interests are at stake.

Furthermore, we are slowly beginning to acknowledge that the increasing autonomy, intelligence, perceptiveness, and empathy of AI, among other features, is moving us away from the idea of the AI robot as a mere tool: some authors argue that we already can develop a closer relationship to social robots than to other objects [1]. Especially as concerns robots –physically embodied AI– the more human they look and act, the more we think they actually have these properties in common with us [4].⁴

At least to the European Parliament the possibility of applying legal personhood to AI does not sound impossible: in its Resolution 2015/2103 (INL), issued on February 16, 2017 [6], the Parliament explicitly suggested the possibility of electronic personhood to the most sophisticated autonomous robots, that interact with people or other robots independently or can make decisions on their own. Electronic personhood is not the same as corporate legal personhood, to be sure, but it is obviously inspired by the corporate form.

In what follows I will discuss a few criticisms that have been raised against the idea of recognizing electronic personhood for AI.

3.1 Criticisms

The idea of ascribing legal personhood to AI has generally not been well received: while the European Parliament seems open to this idea, the European Economic and Social Committee (EESC) [7] explicitly rejects it by arguing that it is opposed to it because of the effects such personhood would have on liability law, possibility of moral hazard and new opportunities for abuse.

In addition, many AI and robotics experts have sent the European Commission an open letter⁵ saying that to grant legal personhood to AI is inappropriate both legally and ethically, and, furthermore, that the justification for this proposition fails to correctly address

³ We will come back to this theory and relation to AI in Section 3.2.

⁴ This is not only a matter of appearance but also one of language: according to Millar and Kerr [5] language matters a lot, because departing from the robot-tool conception we give credit to the uniqueness of AI and acknowledge its social meaning: this is why we have started speaking of a new ontological category in thinking about AI, because the concept of a tools seems to be too limited to capture AI.

the problem of liability. The argument, in particular, is that the analogy between AI and corporations does not hold, since “it implies the existence of human persons behind the legal person to represent and direct it. And this is not the case for a robot.”⁶

Following the same line of thinking is Pagallo [8], who also takes the view that “in the midterm, [we should] skip any hypothesis of granting AI robots full legal personhood,” arguing that legal personhood can be easily substituted with other legal solutions, such as insurance or registries. In addition, drawing on Bryson, Diamantis and Grant [9], Pagallo supports his thesis by deploying what I would call a “double loophole argument,” meaning that an ascription of legal personhood to AI could lead to two risks: humans could use AI to avoid liability, and AI could abuse human rights.

There are many more critical voices, but I focus on the ones just mentioned because they represent the general trend of criticism.

3.2 Response to Criticisms

Let us first look at the EESC argument. What EESC is essentially saying is that if AI were to be recognized as having legal personhood, (a) liability law would be undermined and (b) we would be facing a moral hazard.

To be sure, granting legal personhood to AI could encourage people to try to shift responsibility to AI in the event the use of AI should lead to loss or injury. But corporate legal personhood shows that this shifting of responsibility can be counteracted with the help of the practice known as “piercing the corporate veil,” meaning that in certain situations a court can decide that a corporation cannot be deemed a separate legal person and should instead be treated as the sum of its human shareholders or managers, who could accordingly be held liable for any wrongs the corporation has committed.⁷ This piercing of the corporate veil can be refashioned into a “piercing of the AI veil,” making it possible to get to take action against those who attempt to use AI to do harm or whose use of AI is negligent, whether these individuals are developers or users of the AI in question.

As to the threat that AI as a legal person would pose a moral hazard, none of the ethical guidelines discussed later on in this paper refer to such hazards, nor does the EESC explain what it has in mind by it. There are a few possible interpretations. For instance, this hazard may arise as a side effect of the previously mentioned responsibility-shifting tactic, or this hazard might be related to the prospect of widespread acceptance that AI can be held responsible, since that would upset our moral convictions. Be that as it may, the only possibility for it to become sound is if we decide that the development of intelligent and autonomous AI is

⁵ The open letter is available here: <http://www.robotics-openletter.eu/>.

⁶ The open letter also includes arguments against natural personhood and Anglo-Saxon Trust Model, which we will not deal with in this paper.

⁷ Courts usually take this course of action in cases where third parties need to be protected or when the corporation has been deliberately set up in such a way as to avoid responsibility.

inherently unacceptable and should altogether be banned, as happened with human cloning. There does not seem to be a call for such a ban, and so I will set this argument aside and move on to the next criticism.

As concerns the open letter and the idea that the legal personhood recognized for corporations cannot be applied to AI because there is no one behind AI (as is in case with a corporation), this argument is not sound either: it rests on a theory of corporate personhood that posits the need for “people behind” an enterprise, but this view has never enjoyed any undisputed consensus among legal scholars, many of whom have begun to see the corporation in a different light because of social, economic, informational, technological, and other changes in our culture. Vesting [10] has called it the corporatist transformation of individual liberties, the results of which is that “corporations are conceived as living bodies [...] in which the idea of the individual tends to disappear.” If we pursue the logic behind that transformation, we would see that it would not be inconceivable to extend personhood to AI over the course of its progressive emergence.

Furthermore, the notion that there are always people behind a corporation but not behind AI, isn’t quite correct. For even behind AI we can find people: those who have designed, developed, tested, manufactured, or monitored the AI in question or have used it. Perhaps the whole problem of liability—one of the reasons why AI might need legal personhood—is that there are actually *too many* people behind AI!

As concerns Pagallo, he is not dismissing the question of personhood, nor is he claiming AI personhood to be morally or legally unacceptable. The main focus of this criticism is on the question of when we should bother with issues of AI personhood, and he says that now (or, in his own words, “in the midterm”) is not the time, but he never takes the view that the question will never become relevant.

As to the other legal solutions by which to deal with AI, namely insurance and registries, these devices do not exclude legal personhood: legal persons are more often than not insured, and their registration is mandatory. Registry and insurance are not *alternatives* to AI personhood but *additions* to it [11].

Finally, as concerns the “double loophole argument,” there is no doubt that both abuses and law infringements are possible, and yet we have not eliminated corporate legal personhood (where such examples abound) suggesting that the possibility of abuse is not serious enough to warrant such a move, and that the benefits of ascribing legal personhood outweigh its risks.

The second part of the “double loophole argument” alleges that AI would act to abuse human rights, but it seems too far off to pose a real threat. By the same token, we are still at a stage in the development of AI where we can build ethical and legal

constraints into the design of AI. This is the subject of the ethical guidelines discussed in the next section.

4 Ethical Considerations

As mentioned, the legal personhood of AI needs to be constructed within an ethical framework, setting boundaries that we do not want AI to cross. In the remainder of this section we will look at two ethical guidelines from Canada and Europe so as to see whether the ascription of legal personhood to AI is consistent with them.

4.1 The Montréal Declaration for Responsible Development of AI

In December 2018, the Montréal Declaration for Responsible Development of AI [12] was presented to the public. It contains ten principles for AI researchers to follow.⁸ None of the principles in the Montréal Declaration explicitly or implicitly prohibit the ascription of legal personhood to AI. On the contrary, a couple of the principles may be interpreted in a way as to suggest that autonomous and intelligent machines and software could, as the technology develops, be accorded a form of legal personhood.

The first of these is the principle of solidarity, under which “AIS [Artificial Intelligence Systems] development should not encourage cruel behavior toward robots designed to resemble human beings or non-human animals in appearance or behavior.” This principle could be described as a landmark principle in thinking about the ascription of legal personhood to AI because, for the first time in guidelines of this kind, the non-sentient beings, such as robots, should be protected from cruelty in the same way that humans and animals should be.

The other principle in the Montréal Declaration that is pertinent to the legal personhood of AI is the principle of responsibility, according to which in case of damage or harm caused by AI, which functions properly, its manufacturer or user may not be help responsible.

What does that mean? It means that we could draw a parallel with corporations here: corporation, and not the people, can be held responsible for any harm or damage it may cause, and so has a duty to redress that injury through an outlay of its own resources. The same could apply to AI: it would be required by law to make whole those who have been injured by its operation, using its own resources or an insurance policy to redress the injury. If we accept this principle, we have thereby accepted that AI can be a subject of duties, which is a crucial part of what it means to have legal personhood.

In summary, if we develop the principles contained in the Montréal Declaration, we should see that they support the idea of ascribing legal personhood to AI even if they do not make that proposition explicit in the text. But let us now see what the

⁸ These are the principles of well-being, respect for autonomy, the protection of privacy and intimacy, solidarity, democratic participation, equity, diversity inclusion, prudence, responsibility, and sustainable development.

position of the European Union is and whether it bears any similarities to that of the Canadian ethicists.

4.2 Draft AI Ethics Guidelines for Trustworthy AI

In April 2019, the European Union issued its Draft AI Ethics Guidelines for Trustworthy AI [13], which states that AI should be built bearing in mind respect for human rights and according to the principles of beneficence (do good), non-maleficence (do no harm), autonomy (preserving human agency), justice (being fair), and explicability (operating transparently).

The Guidelines highlights the idea of trustworthy AI, which contains an ethical and a technological component: the ethical component means that a trustworthy AI respects human rights, principles, and values, whereas the technological component requires AI to be technologically robust and reliable. The legal personhood of AI could fall within the limits of this definition: neither of these components is contradicted by the autonomy and intelligence of AI; on the contrary, these properties could work in *favor* of human rights and principles listed above.

Perhaps most relevant to AI as legal persons is the section dedicated to the critical concerns raised by AI, such as the possibility of artificial consciousness: we need to keep an eye on these potential developments, but Guidelines do not propose to ban them altogether.

The same line of reasoning is reflected in requirement of Human Agency and Oversight, according to which less we control AI, “the more extensive testing and stricter governance is required.”

Again, no limits are placed on the autonomy of AI, which means that research on AI and attempts to make it so is not deemed inherently negative. In other words, we need more testing and oversight of autonomous technologies, but it does not mean that they are in themselves harmful or threatening. The Guidelines does not explicitly address the problem of ascribing legal personhood to AI, but the broader framework suggests that they do not take any implied stance against that prospect, either.

In the following section I address what I consider to be among the most important issues involving such an ascription of legal personhood.

5 Issues Involving the Ascription of Legal Personhood to AI

As already hinted at in the previous sections, one of these issues is that we often conflate the legal personhood of AI with the idea that AI and robots should be recognized as human members of society endowed with human rights. But ascribing legal personhood to AI does not necessarily mean conferring human rights on it. Indeed, that corporations in the US are recognized as having constitutional rights is one of the problematic points of

corporate personhood [14], but that is not what happened in Europe. This shows that human rights can be separated from corporate rights and from any AI rights: this is crucial if AI is to be ascribed any rights.

In light of the above, we should have a more serious conversation about legal personhood in general and focus also on the benefits it might bring to all of us. So far we seem to be ignoring that recognizing legal personhood could foster investment in AI, among other benefits, considering that with a recognition of legal personhood also comes the ability to be made whole (by the person so recognized) in the form of damages. AI developers accordingly set up an insurance fund (to which AI itself could contribute with its earnings) and feel less threatened by liability cases that they (the developers) can neither foresee nor control. Consequently the investment in this industry could grow. To be sure, this solution may incentivize the wrong behavior, encouraging AI developers to shift liability to the machines they develop, but if the rules are properly designed, and legal checks are put in place along the way, this kind of possible outcome can be prevented or at least reduced to a minimum. Nor should an ascription of liability to AI be taken to mean that AI developers are completely exempt from liability. What it means, rather, is that the burden of liability is shared between developers and the AI they develop.⁹

Another important question concerns the contents of AI personhood. In particular, there are a few basic rights that will need to be established for AI by considering, for example, whether AI should be able to enter into contracts or report cruel behavior. There is no reason why AI shouldn't have these options. But we need to discuss and define the boundaries of personhood, to which end we could draw on existing frameworks for legal personhood (the right to enter into contracts, a right recognized for humans and corporations alike), but we should also be thinking about AI-specific rights and obligations, such as the right to data and energy [15], upgrades, and maintenance.

Another question still pending, even if it has been addressed by many authors, is that of punishment [16]. We punish corporations by imposing fines, as well as by preventing them from participating in public tenders and dissolving them, among other means. But we still haven't worked out how to punish AI. The possibilities are numerous—from “pulling the plug” on AI to reprogramming it— and the debate is on [17].

6 Closing Remarks

Although the idea of ascribing legal personhood to AI has been widely criticized, there are also many who support the idea [11], [15]. Nor, as discussed, is there a very strong case for denying the legal personhood of AI on moral grounds [9]. Still, if a strong

⁹ On a more practical level, such a change in the paradigm for regulating AI could mean, for example, that we could move sooner, rather than later, in implementing the practice of using personal home assistants for the ill and the elderly, who

according to statistics are increasing in number every year, and are increasingly alone, without anyone in the family being able to act as caretaker

moral argument can not be made today, we cannot be sure that it will continue to be so indefinitely in the future.

There are also several ideas from the field of corporate personhood that we can take inspiration from. For example, when we register a corporation, the corporation is granted a corporate charter by public authorities. This means that a corporation can be made publicly accountable and that its operations should prove beneficial to the public [14], failing which, the corporate charter can be revoked and the corporation would cease to exist. Could we borrow something from this design and adapt it to AI personhood? As mentioned, AI registration seems to be a point everyone agrees on, but then a public registry could also require that a request for registration be justified by explaining what benefits a particular AI would bring into being and how the accountability issue would be dealt with.

There is also another idea from corporate thinking, namely the idea of corporate social responsibility (CSR), which Hamilton [2] describes as “the most powerful attempt to responsabilize the corporation, while at the same time recognizing its personality.” CSR means that corporations owe greater social, economic and environmental responsibilities to society than to its shareholders. Unlike what happened with CSR, which emerged late in corporate history, we could work on a scheme of AI Social Responsibility (AISR) from the very beginning of AI as a legal person separate from its developers and users. For instance, we could implement AISR-by-design, making it mandatory to prove that autonomous and intelligent AI is socially responsible before it is brought into contact with human beings. We could start from the idea of CSR in making the transition toward AISR, using to advantage the fact that AI is capable of AISR-by-design in a way that corporations are not.

The goal of this paper has been to show that legal personhood can reasonably be ascribed to AI. It goes without saying that there will be many obstacles along this path, but none so serious as to prevent legal thought from advancing and finding new ways or reinterpreting the old to deal with contemporary problems.

ACKNOWLEDGMENTS

This paper is part of the project supported by the CONEX programme and has received funding from the Universidad Carlos III de Madrid, the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement N. 600371, el Ministerio de Economía, Industria y Competitividad (COFUND2014-51509), el Ministerio de Educación, Cultura y Deporte (CEI-15-17) and Banco Santander.

REFERENCES

1. K. Darling (2016). Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior towards Robotic Objects. *Robot Law*, eds. R. Calo, A. M. Froomkin and I. Kerr, p. 213–231. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA.
2. S. N. Hamilton (2009). *Impersonations: Troubling the Person in Law and Culture*, p. 18. University of Toronto Press, Toronto, Canada.

3. D. Graver (1999). Personal Bodies: A Corporeal Theory of Corporate Personhood. *University of Chicago Law School Roundtable*, 6(1), 235–250.
4. F. H. Llano Alonso (2018). *Homo Exaltior: Los Límites Ético-Jurídicos del Transhumanismo*. Tirant lo Blanch, Valencia, Spain.
5. J. Millar and I. Kerr (2016). Delegation, Relinquishment, and Responsibility: The Prospect of Expert Robots. *Robot Law*, eds. R. Calo, A. M. Froomkin and I. Kerr, p. 102–127. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA.
6. European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//EN>.
7. Opinion of the European Economic and Social Committee on “Artificial Intelligence – The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society,” own-initiative opinion, 2017/C 288/01, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016IES369&from=EN>.
8. U. Pagallo (2018). Vital, Sophia, and Co. – The Quest for the Legal Personhood of Robots. *Information*, 9(230), 1–11.
9. J. J. Bryson, M. E. Diamantis, T. D. Grant (2017). Of, For and By the People: The Legal Lacuna of Synthetic Persons. *Artificial Intelligence and Law*, 25(3), 273–291.
10. T. Vesting (2018). *Legal Theory and the Media of Law*, p. 552. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA.
11. M. Barrio Andres (2018). Del Derecho de Internet al Derecho de los Robots. *Derecho de los Robots*, ed. M. Barrio Andres, 61–86. Wolster Kluwer, Madrid, Spain.
12. University of Montreal (2018). The Montreal Declaration for Responsible Development of AI, <https://www.montrealdeclaration-responsibleai.com/the-declaration>.
13. EU Commission’s High Level Expert Group on Artificial Intelligence (2019). Draft AI Ethic Guidelines for Trustworthy AI, <https://ec.europa.eu/futurium/en/ai-alliance-consultation>.
14. J. D. Clements. *Corporations are not People: Reclaiming Democracy from Big Money and Global Corporations* (2nd. ed.). Berrett-Koehler Publishers, San Francisco, CA, USA.
15. W. Wallach and C. Allen (2008). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press, New York, NY, USA.
16. U. Pagallo and S. Quattrocchio (2018). The Impact of AI on Criminal Law, and Its Two Fold Procedures. *Research Handbook on the Law of Artificial Intelligence*, ed. W. Barfield and U. Pagallo, 385–409. Edward Elgar Pub, Cheltenham, UK.
17. G. Hallevy (2018). Dangerous Robots – Artificial Intelligence vs. Human Intelligence, Hallevy, Prof. Gabriel, *Dangerous Robots – Artificial Intelligence vs. Human Intelligence* (February 11, 2018), <https://ssrn.com/abstract=3121905> or <http://dx.doi.org/10.2139/ssrn.3121905>.