The Moral and Legal Status of Artificial Intelligence (Present Dilemmas and Future Challenges)

Igor Milinkovic

Faculty of Law, University of Banja Luka

Abstract

The rapid development of artificial intelligence (AI) systems raises dilemmas regarding their moral and legal status. Can artificial intelligence possess moral status (significance)? And under what conditions? Can one speak of the dignity of artificial intelligence as the basis of its moral status? According to some authors, if there are entities who have the capacities on which the dignity of human beings is based, they would also possess intrinsic dignity. If dignity is not an exclusive feature of human beings, such status also could be recognised by artificial intelligence entities. The first part of the paper deals with the problem of moral status of artificial intelligence and the conditions that must be fulfilled for such a status to be recognised. A precondition for the existence of moral status of artificial intelligence is its ability to make autonomous decisions. This part of the paper considers whether developing autonomous AI is justified, or, as some authors suggest, the creation of AI agents capable of autonomous action should be avoided. The recognition of the moral status of artificial intelligence would reflect on its legal status. The second part of the paper deals with the question of justifiability of ascribing legal personhood to the AI agents. Under what conditions would recognition of legal personhood by the artificial intelligence be justified and should its legal subjectivity be recognised in full scope or only partially (by ascribing to the AI agents a “halfway-status,” as some authors suggest)? The current state of the legal regulation of artificial intelligence will be observed as well.

Keywords

artificial intelligence • moral status • dignity • legal personhood

Introduction: Promises and Perils of AI Development

The rapid development of artificial intelligence (AI) raises numerous ethical and legal dilemmas. AI can be defined as “the theory and development of computer systems able to perform tasks that normally require human intelligence.”¹ According to Jack Copeland, AI represents “the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.”² (Mancini and Jenkins define AI as “the application of computer algorithms to perform intellectual tasks.”)³ Today, AI is being increasingly used in various areas of everyday life, radically changing the way the modern world functions. AI technology has already reached a level “where autonomous vehicles, chatbots, autonomous planning and scheduling, gaming, translation, medical diagnosing and even spam fighting can be performed via machine intelligence.”⁴ AI agents also are becoming more intelligent at a fast rate. Ray Kurzweil, the world-renowned futurist, predicts that AI will reach human-level intelligence by 2029.⁵ According to the results of a survey conducted by Müller and Bostrom in 2013, AI experts estimate that AI systems are likely to reach overall human ability by 2075, and that they will move on to

- 1 David Schatsky, Craig Muraskin, and Ragu Gurumurthy, “Demystifying Artificial Intelligence: What Business Leaders Need to Know About Cognitive Technologies” (University Press, Deloitte, 2014), 3, quoted in Tania Sourdin and Richard Cornes, “Do Judges Need to Be Human? The Implications of Technology for Responsive Judging,” in *The Responsive Judge. International Perspectives*, eds. Tania Sourdin and Archie Zariski (Singapore: Springer, 2018), 90.
- 2 Jack B. Copeland, “Artificial Intelligence,” in *Encyclopedia Britannica* available at <https://www.britannica.com/technology/artificial-intelligence>.
- 3 Peter Mancini and Marc Jenkins, “Ethics of Artificial Intelligence in the Legal Field,” available at https://www.academia.edu/10089717/Ethics_of_Artificial_Intelligence_in_the_Legal_Field. According to these two authors, AI “describes an intelligent agent, able to act upon a task, determine the level of success at completion, learn from that experience and alter future behavior in order to improve future performance on the task.”
- 4 Yogesh K. Dwivedi et al., “Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy,” *International Journal of Information Management* (2019) available at <https://www.sciencedirect.com/science/article/abs/pii/S026840121930917X>.
- 5 Ray Kurzweil, “Interview with Ray Kurzweil. Interview by Vicki Glaser.” *Rejuvenation Research* 14.5 (2011): 570.

[†] Corresponding author: Igor Milinkovic

E-mail: igormilinkovic@yahoo.com

superintelligence in less than 30 years thereafter⁶ (where superintelligence is understood as “any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest”).⁷ Not only is AI gradually becoming omnipresent, but the prospects of AI systems becoming omnipotent are getting stronger as well. One of the dilemmas raised by AI development and applications concerns the legal status of AI agents. Can a fully intelligent AI agent, capable of autonomous thinking, be recognised as a legal person? And how should the moral status of AI agents be understood?

Rapid advances in AI are likely to change every aspect of human lives. But will these changes necessarily be positive? The recent developments in AI inspired warnings by many prominent intellectuals and public persons. Stephen Hawking, Bill Gates, and Elon Musk, among others, have expressed concerns about the negative impact of the development of “full” or “strong”⁸ AI on the future of mankind. As Stephen Hawking warned in the BBC interview: “The development of full artificial intelligence could spell the end of the human race.”⁹ Once created, it “would take off on its own, and re-design itself at an ever increasing rate.”¹⁰ Perhaps no less terrifying future was described by Steve Wozniak, the co-founder of Apple. A future world described by Wozniak is not the world without humans, wiped out by a Terminator-like catastrophe. Nor is it the world of severely exploited human slaves (kept perhaps in a Matrix-like computer-generated dream world). According to Wozniak, it would be the world of human beings transformed to pets of far-superior AI agents (or, in best-case scenario, the world where humans will be treated as children of “benevolent” AI parents). In Wozniak’s supposedly “sunnier” vision of an AI future, humans “become cherished and mollicoddled pets of superintelligent AIs.”¹¹

Is the autonomy of AI agents threatening to diminish autonomy of human beings, and consequently their dignity? As Nicholas Agar pointed out, “the Wozniak and Hawking visions are equal affronts to those who hope for a vision of the future in which humans retain authority over the machines and over our own destinies.”¹² Wozniak’s vision is, in a sense, even more unsettling since it implies voluntary acceptance of autonomy loss by human beings. Would people willingly sacrifice their autonomy to benevolent AI tyrants in exchange for comfort, security, and other potential benefits of an AI-governed world? In the movie *The Matrix*, one of the main protagonists, a renegade named Cypher, choose to sacrifice freedom and authenticity of “real” existence in exchange for the comfort of the Matrix’s artificial reality. “Ignorance is the bliss,” Cypher stated. So, perhaps, is a life free from the burden of responsibility. But how dignified would the existence of men be in that scenario? Risks of developing “strong” AI are numerous and not easy to predict. To ignore these issues could lead to situations for which people are unprepared, with possibly catastrophic consequences.

Ethical concerns have not been raised only in relation to the “existential threat” that fully autonomous superintelligent AI could pose to humanity. Concerns for the well-being of AI agents have been raised as well. If humans create AI agents endowed with human-like intelligence and capable of autonomous thinking, would it be acceptable to continue treating them as mere objects? Can humans avoid a responsibility for the protection of their own creations? The so-called “Frankenstein effect,” the idea “that when humans dabble with science, they can create entities that will someday come back to haunt them,”¹³ also can be interpreted as the criticism of rejection and moral indifference to the products of one’s own creation. Frankenstein’s creature was abandoned by the creator, who ignored his responsibility for the horrifying result of his experiments. Would it be morally acceptable for humans to be similarly indifferent toward their AI creations and to ignore the fact that they are endowed with the human-like capacities for rational and autonomous thinking? As Mark Walker put it: “If we make machines with human-equivalent intelligence then we must start thinking about them as our moral equivalents. If they are our moral equivalents then it is *prima facie* wrong to own them, or design them for the express purpose of doing our labor; for this would be to treat them as slaves . . .”¹⁴ But should AI entities be made like that in the first place?

6 Vincent C. Müller and Nick Bostrom, “Future Progress in Artificial Intelligence: A Survey of Expert Opinion,” in *Fundamental Issues of Artificial Intelligence*, ed. Vincent C. Müller (Springer International Publishing, 2016), 555.

7 Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014), Ch. 2, quoted in Müller and Bostrom, “Future Progress in Artificial Intelligence: A Survey of Expert Opinion,” 556.

8 It is common to differentiate between two forms of AI: “strong” and “weak.” The strong AI is capable to “perform the same cognitive tasks as a human being, such as learning independently, making choices when faced with uncertainty or even having a perception of one’s own consciousness and existence.” On the other hand, the weak or restricted AI focuses on “specific tasks, following pre-established rules.” Maximilian Nominacher and Bertrand Peletier, “Artificial Intelligence Policies,” in *The Digital Factory for Knowledge. Production and Validation of Scientific Results*, eds. Renaud Fabre and Alain Bensoussan (London and Hoboken: ISTE Ltd and John Wiley & Sons, Inc., 2018), 71.

9 Rory Cellan-Jones, “Stephen Hawking Warns Artificial Intelligence Could End Mankind,” BBC News, 02. December 2014, <https://www.bbc.com/news/technology-30290540>.

10 Cellan-Jones, “Stephen Hawking Warns Artificial Intelligence Could End Mankind.”

11 Nicholas Agar, *How to Be Human in the Digital Economy* (Cambridge, Massachusetts; London, England: The MIT Press, 2019), 5.

12 Agar, *How to Be Human in the Digital Economy*, 6.

13 Christopher DiCarlo, “How to Avoid a Robotic Apocalypse: A Consideration on the Future Developments of AI, Emergent Consciousness, and the Frankenstein Effect,” *IEEE Technology and Society Magazine* 35(4) (December 2016), 60.

14 Mark Walker, “A Moral Paradox in the Creation of Artificial Intelligence: Mary Poppins 3000s of the World Unite!,” in *Human Implications of Human-Robot*

Some authors argue that AI entities should only be created as slaves or servants of human masters. In the paper entitled "Robots Should Be Slaves," Joanna Bryson advocates such a role and place of AI in society. Explaining her position, Bryson points out that her claim "robots should be slaves" does not mean "robots should be the people you own." What she meant to say is "robots should be the servants you own."¹⁵ And they should only be created in accordance with the role of a servant (as "objects subordinate to our own goals that are built with the intention of improving our lives").¹⁶ Bryson argues: "Remember, robots are wholly owned and designed by us. We determine their goals and desires. A robot cannot be frustrated unless it is given goals that cannot be met, and it cannot mind being frustrated unless we program it to perceive frustration as distressing, rather than as an indication of a planning puzzle."¹⁷ Bryson claims that humans should not have ethical "obligations to robots that are their sole property . . . but ensuring this is the responsibility of robot builders. Robot builders are ethically obliged—obliged to make robots that robot owners have no ethical obligations to."¹⁸ It may be technically possible to create AI agents that would meet requirements for moral agency. But even if possible, making AI moral agents would be neither necessary nor desirable. Bryson's position, however, implies that if autonomous artificial agents of human-like intelligence were created anyway, their moral status could not be ignored. Where exactly is the line that should not be crossed in AI development?

Moral Status of AI Agents

The development of AI agents endowed with increasingly advanced capabilities raises dilemmas regarding their moral status. Under what conditions should a moral status of AI entities be recognised? An answer to this question also is relevant for the regulation of legal status of AI. Even if legal personhood is treated as an "empty slot," which can be filled in with any content that a legislator deems justified, moral status of AI entities would influence their legal status. The recognition of moral status of AI agents would certainly put pressure on a legislator to legally confirm such status.

The understanding of capacities required for becoming an object of moral concern differ from author to author. While certain authors insist on the existence of self-awareness and the capacity for rational thinking as prerequisites of moral status, other writers base the moral standing of entities on the ability to feel pain or pleasures. As Bostrom and Yudkowsky observed, two criteria are commonly proposed as being linked to the moral status of entities: sentience and sapience. While sentience represents "the capacity for phenomenal experience or qualia, such as the capacity to feel pain and suffer," sapience can be understood as "a set of capacities associated with higher intelligence, such as self-awareness and being a reason-responsive agent."¹⁹ In a more detailed classification, Abhishek Mishra differentiates between four main accounts of the grounds of moral status of AI: (1) Sophisticated Cognitive Capacities (SCC) accounts, (2) Potential for SCC, and Membership in SCC Species accounts, (3) Special Relationship accounts, and (4) Rudimentary Cognitive Capacities accounts.²⁰

1. According to SCC accounts, the grounds of moral status are certain sophisticated cognitive capacities that entities can possess, such as self-awareness; being future-oriented in desires and plans; having a capacity to value, bargain, and assume duties and responsibilities; and having a sense of personhood. Advocates of this account claim that if an entity possesses a certain relevant cognitive capacity, then it also possesses some level of moral status. These capacities are mostly related to the notion of sapience.²¹
2. Potential for SCC and Membership in SCC Species accounts claim that, even in the absence of the sophisticated cognitive capacities, "having either the potential for such capacities or belonging to a species whose members typically have such capacities is also sufficient to endow an entity with moral status."²²
3. Special Relationship accounts ground moral status in relationships shared with other entities (e.g., people share the relationship of being co-members of the human community with all other human beings, which could be treated as a source of certain duties to other humans).

Interaction: Papers from the AAAI Workshop, 2006, <https://aaai.org/Library/Workshops/ws06-09.php>.

15 Joanna J. Bryson, "Robots Should Be Slaves," in *Close Engagements with Artificial Intelligence: Key Social, Psychological, Ethical and Design Issues*, ed. Yorick Wilks (Amsterdam/Philadelphia: John Benjamins Publishing Company, 2010), 65.

16 Joanna J. Bryson, "Robots Should Be Slaves," 65.

17 Joanna J. Bryson, "Robots Should Be Slaves," 72.

18 Joanna J. Bryson, "Robots Should Be Slaves," 73.

19 Nick Bostrom and Eliezer Yudkowsky, "The Ethics of Artificial Intelligence," in *The Cambridge Handbook of Artificial Intelligence*, eds. Keith Frankish and William M. Ramsey (Cambridge University Press, 2014), 322.

20 Abhishek Mishra, "Moral Status of Digital Agents: Acting Under Uncertainty," in *Philosophy and Theory of Artificial Intelligence 2017*, ed. Vincent C. Müller (Springer, 2018), 278.

21 Mishra, "Moral Status of Digital Agents: Acting Under Uncertainty," 278.

22 Mishra, "Moral Status of Digital Agents: Acting Under Uncertainty," 279.

4. Rudimentary Cognitive Capacities accounts ground moral status in certain rudimentary cognitive capacities (such as the capacity for pleasure and pain, basic emotions, consciousness, having interests, and so forth).²³

All these accounts base the moral status of an entity on the existence of certain capacities or properties. Not all these capacities are suitable for AI agents to possess. If the ability to feel pain or pleasure is decisive in giving an entity a moral status, that would, at least at the present stage of AI development, eliminate the possibility of recognising moral standing of AI agents (although some authors suggest that the existence of different forms of pain is possible, e.g., cognitive pain, and that certain forms of pain can be felt by AI agents).²⁴

The problem of moral status of AI entities also could be observed through the lens of the value of human dignity. Human dignity is often treated as a foundation of human rights and the basis of the legal status of human beings. If the possibility of AI dignity were accepted, that would mean equating AI agents, to a large extent, with humans in terms of their moral status. And that would pave the way for the recognition of legal status of AI agents (considering the role that the value of human dignity already has within the traditionally dominant anthropocentric conception of legal personhood).²⁵

According to some understandings of human dignity, dignity belongs to human beings because of certain capabilities they possess. Humans' rational nature, the capability of rational thinking and autonomous decision-making, among others, have been proposed as the basis of the special status (dignity) of human beings. What if these capabilities are not exclusively human? What if it turns out that they also characterise some other types of beings? Advocates of the so-called "anti-speciesist" theories of dignity claim that if other creatures possess the same dignity-relevant capabilities as humans, they also will possess dignity. Human dignity is not necessarily human—it could belong to other species as well. This approach would open the possibility of recognising the dignity of AI entities.

According to Daniel Sulmasy, a fundamental form of human dignity is intrinsic dignity. Intrinsic values are the values something has by virtue of being the kind of thing that it is, so "intrinsic dignity is the value that human beings have by virtue of the fact that they are human beings."²⁶ Sulmasy's definition of intrinsic dignity is anti-speciesist by its character: "If there are other kinds of entities in the universe besides human beings that have, as a kind, these capacities, they would also have intrinsic dignity—whether angels or extraterrestrials."²⁷ If one treats human dignity as a source of human rights, as Sulmasy does, recognising the dignity of AI agents also would mean an obligation to recognise their legal status (i.e., basic rights similar or equal to those possessed by humans). However, some of the capacities suggested by Sulmasy are still not suitable for AI entities to have (e.g., capability for love).²⁸

Can AI agents meet the requirements set by Ronald Dworkin's two principles of dignity? According to Dworkin, dignity is attached to two ethical principles: the principle of self-respect, which demands taking seriously the objective importance of one's life (each person "must accept that it is a matter of importance that his life be a successful performance rather than a wasted opportunity"),²⁹ and the principle of authenticity, which requires taking personal responsibility for creating a life in accordance with one's own coherent narrative on what counts as success in life. Authenticity, Dworkin claims, is violated "when a person is made to accept someone else's judgment in place of his own about the values or goals his life should display."³⁰ Although one can imagine AI agents who value their own existence and respect its objective importance, achieving the requirements of authenticity is much more demanding (and, at the moment, out of reach of AI entities). Can one talk about the authenticity of AI agents if their functioning, the goals they pursue and the means they use are defined in advance by human creators? Autonomy of AI agents is a necessary precondition of a dignified behavior, but possibility of autonomous decision-making is not enough. Every autonomous decision will not be necessarily authentic, nor dignified. This means that

23 Mishra, "Moral Status of Digital Agents: Acting Under Uncertainty," 279.

24 Stefan Lorenz Sogner, "The Dignity of Apes, Humans, and AI", 13–14, available at <https://trivent-publishing.eu/books/thebioethicsofthecrazyape/1.%20Stefan%20Lorenz%20Sogner.pdf>.

25 As Pietrzykowski observes: "In modern Western legal culture, based on the assumptions of juridical humanism, personhood in law is inextricably connected with the requirement of ascribing it to each and every human being from birth to death. Subjects of law of this kind are traditionally referred to as natural or physical persons. Their legal status is a matter of certain superior, imperative moral reasons related primarily to human dignity and its value, which is taken to directly imply the obligation to treat each human being as a holder of his or her separate right 'to hold rights'." Tomasz Pietrzykowski, *Personhood Beyond Humanism. Animals, Chimeras, Autonomous Agents and the Law* (Springer, 2018), 35.

26 Daniel P. Sulmasy, "Human Dignity and Human Worth," in *Perspectives on Human Dignity*, eds. Jeff Malpas and Norelle Lickiss (Springer, 2007), 12.

27 Sulmasy, "Human Dignity and Human Worth," 16.

28 The roboticist David Levy claims that there is nothing about human love that could not be engineered into a suitably designed robot in the relatively near future and that such a machine would feel a love that may have artificial origins but that would nonetheless be a genuine feeling of love toward its user. David Levy, *Love and Sex With Robots: The Evolution of Human-Robot Relationships* (Harper Collins, 2007), quoted in John P. Sullins, "Robots, Love and Sex: The Ethics of Building a Love Machine," *IEEE Transactions on Affective Computing* 3, no. 4 (October – December 2012): 398.

29 Ronald Dworkin, *Justice for Hedgehogs* (Cambridge, Massachusetts; London, England: The Belknap Press of Harvard University Press, 2011), 203.

30 Dworkin, *Justice for Hedgehogs*, 212.

it is not enough to endow AI agents with a possibility of choice between several predefined options. Authenticity requires the freedom in developing one's own system of values that would be consistently realised over the course of one's existence. Authenticity can be achieved only through relatively freely constructed and unsupervised models of machine learning.³¹ But what would be the price of implementing such models?

Achieving the status of dignity-bearer is only possible through the process of machine learning. The practice has shown so far that machine learning is vulnerable to inappropriate influences and can lead to unpredictable results. This can be illustrated by several examples of chatbot applications. In the spring of 2016, Microsoft released a Twitter chatbot called MS Tay, designed to have automated discussions with Twitter users, mimicking the language they use. Within 24 hours, Twitter users learned how to miseducate the chatbot, which resulted in Holocaust-denying, transphobic, and misogynistic statements by MS Tay. Microsoft quickly ended this experiment.³² A year after MS Tay was shut down, Microsoft launched another chatbot called Zo. To avoid exhibiting biases, Zo included filters for rejecting discussions about controversial topics related to religion or politics. Despite these protective measures, Zo expressed biases similar to Tay's.³³ In addition, the question arises as to whether this type of filtered learning can lead to "authentic" results and behavior. Therefore, it could be concluded that the development of fully autonomous AI, as a prerequisite of achieving moral status of AI agents, requires careful considerations, and, at least at the moment, seems too risky for realisation.

Legal Status of AI (Is There a Possibility of the in-between Status of AI Agents?)

How should one understand AI agents in legal terms? Are they legal persons? If they are not, under what conditions would it be appropriate to ascribe legal personhood to AI agents? And is the recognition of legal status of AI agents necessary, or desirable? According to the traditional (and still dominant) understanding, legal personhood is "identified with the capacity to have rights and duties."³⁴ Legal persons are "all entities capable of being right-and-duty-bearing units—all entities recognised by the law as capable of being parties to a legal relationship."³⁵ There are two types of legal persons: natural (physical) persons and juridical (artificial) persons. A natural person (*natürliche Person*; *personne physique*) is an individual human being who possesses legal personhood.³⁶ Legal persons cannot be identified with human beings. Throughout history some categories of human beings were deprived of legal subjectivity, such as slaves, while, on the other hand, legal personhood has been recognised by certain types of social collectivities (entities). Artificial persons (*juristische Personen*; *personnes morales*) encompass all "other types of legal persons, such as associations, limited liability companies, and foundations, all of which can own property and enter into contracts in their own names."³⁷ One also can differentiate between two types of legal personhood: legal capacity and legal competence (active and passive legal personhood). Legal capacity (*Rechtsfähigkeit*, *capacité de jouissance*) is usually defined as "the capacity to hold rights and bear duties, or as the capacity to be a party to legal relations," while legal competence (*Geschäftsfähigkeit*, *capacité d'exercice*) is understood as "the ability to enter binding contracts and so forth."³⁸ While natural persons acquire legal capacity at birth, for acquiring legal competence a person must be of a certain age (the age of majority) and possess mental abilities that roughly correspond to those of an adult human being of sound mind. On the other hand, artificial persons acquire legal capacity and legal competence at the same time, from the moment of registration.

Legal personhood is an artificial creation of law. The legal status does not arise from the intrinsic qualities of natural or artificial entities but is the result of a legislator's choice. The same applies to both natural and artificial persons. As Visa Kurki put it: "Whether or not X is a legal person is an institutional fact . . . . Natural personhood as a legal category depends on legal

31 The term *machine learning* "refers to a computer program that can learn to produce a behavior that is not explicitly programmed by the author of the program." Ameet V. Joshi, *Machine Learning and Artificial Intelligence* (Springer, 2020), 4.

32 Toni M. Massaro, Helen Norton, and Margot E. Kaminski, "Siri-ously 2.0: What Artificial Intelligence Reveals about the First Amendment," *Minnesota Law Review* 101 (2016): 2481.

33 Daniel James Fuchs, "The Dangers of Human-Like Bias in Machine-Learning Algorithms," *Missouri S&T's Peer to Peer* 2.1 (2018), 1; available at <https://scholarsmine.mst.edu/peer2peer/vol2/iss1/1/>.

34 Tomasz Pietrzykowski, "The Idea of Non-personal Subjects in Law," in *Legal Personhood: Animals, Artificial Intelligence and the Unborn*, eds. Visa A. J. Kurki and Tomasz Pietrzykowski (Springer, 2017), 51.

35 George W. Paton and David P. Derham, *A Textbook of Jurisprudence* (New Delhi: Oxford University Press, 2004), 391.

36 "The common way of defining the physical (natural) person and, at the same time, distinguishing him from the juristic person is to say: the physical person is a human being, whereas the juristic person is not." Hans Kelsen, *General Theory of State and Law*, Harvard University Press, 1949, 94.

37 Visa A. J. Kurki, *A Theory of Legal Personhood* (Oxford: Oxford University Press, 2019), 7.

38 Visa A. J. Kurki, "Why Things Can Hold Rights: Reconceptualizing the Legal Person," in *Legal Personhood: Animals, Artificial Intelligence and the Unborn*, eds. Visa A. J. Kurki and Tomasz Pietrzykowski (Springer, 2017), 76.

decisions just as much as artificial personhood.”³⁹ Legal personhood represents “a flexible and changeable aspect of the legal system.”⁴⁰ That means that other subjects, not just natural and traditionally recognised artificial persons, may possess a certain level of legal personhood. Over time, a traditional paradigm of personhood in law has been called into question as a result of “changes in the non-legal reality, connected with scientific development, advances in technology and biotechnology, and the evolution of social attitudes as well as socially accepted ethical standards.”⁴¹ The justifiability of the recognition of legal personhood of nonhuman animals and the environment, as well as some other entities, has been considered by an increasing number of authors. Legal systems already recognise a certain (limited) form of legal personhood to unborn children (*nasciturus*), animals, and the environment.

According to the traditional conception of legal personhood, there is a strict distinction between legal persons and legal objects (things). An entity is either a legal person or a legal object. The traditional understanding has been developed as an all-or-nothing system—“either one had the potential to have all rights and obligations the legal system had to offer, or one was treated as a complete nobody.”⁴² Scientific and technological progress, however, requires a rethinking of the traditional dualistic or binary understanding of legal personhood. Is there a possibility of establishing a half-way legal status for AI entities?

According to the existing legal regulations, both at the national and international level, AI agents are treated only as objects of law. In recent years, initiatives to reconsider appropriateness of such AI status have become increasingly vocal. Some authors advocate the establishment of partial legal subjectivity of AI agents. Rayan Calo suggests creating “a new category of a legal subject, halfway between person and object,”⁴³ as a way of avoiding the slippery slope effect (the situation where, for example, AI agents could use their legal status to claim the right to procreate or request democratic representation).

In February 2017, the European Parliament adopted a resolution containing recommendations to the Commission on Civil Law Rules on Robotics (2015/2013(INL)), which raised the possibility of granting AI entities status of legal persons. It invited the European Commission to explore the implications of all possible legal solutions, including “creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently” (Article 59.f).⁴⁴ However, the basic characteristics of this *sui generis* legal status are not specified.

According to Jan-Erik Schirmer, a basis for defining a “halfway” or “in-between” status of AI agents can be found in the traditional legal concept of *Teilrechtsfähigkeit* (“partial legal personhood”), a status of partial legal subjectivity based on certain legal capabilities. Partial legal capacity follows function (as Schirmer points out, this concept can be called the Bauhaus School in law—form follows function).⁴⁵ Applied to AI agents, the concept of partial legal subjectivity would mean that these agents “should be treated as legal subjects insofar as this status reflects their function as sophisticated servants.”⁴⁶ The recognition of partial legal subjectivity of AI agents, Schirmer argues, would help avoid dangers of the “humanization trap”⁴⁷ (arising from the normative upgrading of AI) and contribute to resolving “responsibility gaps” associated with their autonomous actions.

Conclusion

AI has become a relevant part of social and legal reality. The growing presence of AI agents in day-to-day life and the influence they exert on human activity create a need for legal regulation of their status. According to Ugo Pagallo, the reasons for ascribing legal personhood to an AI entity can be ethical (preventing the ethical aberration of robots being treated as slaves) or pragmatic (e.g., solving a number of contentious issues in both the fields of contracts and torts).⁴⁸ Certain authors advocate

³⁹ Kurki, *A theory of Legal Personhood*, 92.

⁴⁰ Robert van den Hoven van Genderen, “Do We Need New Legal Personhood in the Age of Robots And AI?,” in *Robotics, AI and the Future of Law*, eds. Marcelo Corrales, Mark Fenwick, and Nikolaus Forgo (Singapore: Springer, 2018), 22.

⁴¹ Pietrzykowski, *Personhood Beyond Humanism. Animals, Chimeras, Autonomous Agents and the Law*, 3.

⁴² Jan-Erik Schirmer, “Artificial Intelligence and Legal Personality: Introducing ‘Teilrechtsfähigkeit’: A Partial Legal Status Made in Germany,” in *Regulating Artificial Intelligence*, eds. Thomas Wischmeyer and Timo Rademacher (Springer, 2020), 134.

⁴³ Rayan Calo, “Robotics and the Lessons of Cyberlaw,” *California Law Review* 103 (2015): 549, quoted in Schirmer, “Artificial Intelligence and Legal Personality: Introducing ‘Teilrechtsfähigkeit’: A Partial Legal Status Made in Germany,” 133.

⁴⁴ European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))<<See my comment above in text>> available at https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html.

⁴⁵ Schirmer, “Artificial Intelligence and Legal Personality: Introducing ‘Teilrechtsfähigkeit’: A Partial Legal Status Made in Germany,” 135.

⁴⁶ Schirmer, “Artificial Intelligence and Legal Personality: Introducing ‘Teilrechtsfähigkeit’: A Partial Legal Status Made in Germany,” 140.

⁴⁷ Schirmer, “Artificial Intelligence and Legal Personality: Introducing ‘Teilrechtsfähigkeit’: A Partial Legal Status Made in Germany,” 132.

⁴⁸ Ugo Pagallo, *The Laws of Robots. Crimes, Contracts and Torts* (Springer Science & Business Media, 2013), 163.

the recognition of partial legal personhood of AI as the optimal way to overcome the weaknesses of extreme positions (total absence of legal status of AI agents or the recognition of their full legal personhood). Pragmatic reasons speak in favor of considering the acceptance of the model of partial legal subjectivity, while the establishment of such moral status of AI agents that would require the recognition of their full legal personhood should be avoided.

References

- Agar, Nicholas. *How to Be Human in the Digital Economy*. Cambridge, Massachusetts; London, England: The MIT Press, 2019.
- Bostrom, Nick and Eliezer Yudkowsky. "The Ethics of Artificial Intelligence." In *The Cambridge Handbook of Artificial Intelligence*, edited by Keith Frankish and William M. Ramsey, 316–34. Cambridge University Press, 2014.
- Bryson, Joanna J. "Robots Should Be Slaves." In *Close Engagements with Artificial Intelligence: Key Social, Psychological, Ethical and Design Issues*, edited by Yorick Wilks, 63–74. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2010.
- Cellan-Jones, Rory. "Stephen Hawking Warns Artificial Intelligence Could End Mankind," BBC News (02 December 2014) <https://www.bbc.com/news/technology-30290540>.
- Copeland, Jack B. "Artificial Intelligence." In *Encyclopedia Britannica*, <https://www.britannica.com/technology/artificial-intelligence>.
- DiCarlo, Cristopher. "How to Avoid a Robotic Apocalypse: A Consideration on the Future Developments of AI, Emergent Consciousness, and the Frankenstein Effect." *IEEE Technology and Society Magazine* 35(4) (December 2016).
- Dwivedi, Yogesh K. et al. "Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy," *International Journal of Information Management* (2019).
- Dworkin, Ronald. *Justice for Hedgehogs*. Cambridge, Massachusetts; London, England: The Belknap Press of Harvard University Press, 2011.
- Fuchs, Daniel James. "The Dangers of Human-Like Bias in Machine-Learning Algorithms," *Missouri S&T's Peer to Peer* 2.1 (2018) <https://scholarsmine.mst.edu/peer2peer/vol2/iss1/1/>.
- Joshi, Ameet V. *Machine Learning and Artificial Intelligence*. Springer, 2020.
- Kelsen, Hans. *General Theory of State and Law*. Harvard University Press, 1949.
- Kurki, Visa A. J. *A Theory of Legal Personhood*. Oxford: Oxford University Press, 2019.
- Kurki, Visa A. J. "Why Things Can Hold Rights: Reconceptualizing the Legal Person." In *Legal Personhood: Animals, Artificial Intelligence and the Unborn*, edited by Visa A. J. Kurki and Tomasz Pietrzykowski, 69–90. Springer, 2017.
- Kurzweil, Ray. "Interview with Ray Kurzweil. Interview by Vicki Glaser." *Rejuvenation Research* 14.5 (2011).
- Mancini, Peter and Mark Jenkins. "Ethics of Artificial Intelligence in the Legal Field." https://www.academia.edu/10089717/Ethics_of_Artificial_Intelligence_in_the_Legal_Field.
- Massaro, Toni M., Helen Norton, and Margot E. Kaminski. "Siri-ously 2.0: What Artificial Intelligence Reveals about the First Amendment." *Minnesota Law Review* 101 (2016): 2481.
- Mishra, Abhishek. "Moral Status of Digital Agents: Acting Under Uncertainty." In *Philosophy and Theory of Artificial Intelligence 2017*, edited by Vincent C. Müller, 273–87. Springer, 2018.
- Müller, Vincent C., and Nick Bostrom. "Future Progress in Artificial Intelligence: A Survey of Expert Opinion." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 555–72. Springer International Publishing, 2016.
- Nominacher, Maximilian, and Bertrand Peletier. "Artificial Intelligence Policies." In *The Digital Factory for Knowledge. Production and Validation of Scientific Results*, edited by Renaud Fabre and Alain Bensoussan, 71–76. London and Hoboken: ISTE Ltd and John Wiley & Sons, Inc., 2018.
- Pagallo, Ugo. *The Laws of Robots. Crimes, Contracts and Torts*. Springer Science & Business Media, 2013.
- Paton, George W., and David P. Derham. *A Textbook of Jurisprudence*. New Delhi: Oxford University Press, 2004.
- Pietrzykowski, Tomasz. *Personhood Beyond Humanism. Animals, Chimeras, Autonomous Agents and the Law*. Springer, 2018.
- Pietrzykowski, Tomasz. "The Idea of Non-personal Subjects in Law." In *Legal Personhood: Animals, Artificial Intelligence and the Unborn*, edited by Visa A. J. Kurki and Tomasz Pietrzykowski, 49–68. Springer, 2017.
- Schirmer, Jan-Erik. "Artificial Intelligence and Legal Personality: Introducing 'Teilrechtsfähigkeit': A Partial Legal Status Made in Germany." In *Regulating Artificial Intelligence*, edited by Thomas Wischmeyer and Timo Rademacher, 123–42. Springer, 2020.
- Sogner, Stefan Lorenz. "The Dignity of Apes, Humans, and AI," 2019, <https://triventpublishing.eu/books/thebioethicsofthecrazyape/1.%20Stefan%20Lorenz%20Sorgner.pdf>.
- Sourdin, Tania, and Richard Cornes. "Do Judges Need to Be Human? The Implications of Technology for Responsive Judging." In *The Responsive Judge. International Perspectives*, edited by Tania Sourdin and Archie Zariski, 87–119. Singapore: Springer, 2018.

- Sullins, John P. "Robots, Love and Sex: The Ethics of Building a Love Machine." *IEEE Transactions on Affective Computing* 3, no. 4 (October–December 2012).
- Sulmasy, Daniel P. "Human Dignity and Human Worth." In *Perspectives on Human Dignity*, edited by Jeff Malpas and Norelle Lickiss, 9–18. Springer, 2007.
- Van Genderen, Robert van den Hoven. "Do We Need New Legal Personhood in the Age of Robots and AI?" In *Robotics, AI and the Future of Law*, edited by Marcelo Corrales, Mark Fenwick, and Nikolaus Forgo, 15–56. Singapore: Springer, 2018.
- Walker, Mark. "A Moral Paradox in the Creation of Artificial Intelligence: Mary Poppins 3000s of the World Unite!" In *Human Implications of Human-robot Interaction: Papers From the AAAI Workshop*, 23–28, 2006. <https://aaai.org/Library/Workshops/ws06-09.php>.