Scientia Moralitas International Journal of Multidisciplinary Research
ISSN 2472-5331 (Print) | ISSN 2472-5358 (Online) | Vol. 4, No. 1, 2019
DOI: 10.5281/zenodo.3355697

The Need for an International Treaty for AI from the Perspective of Human Rights

Themistoklis Tzimas, PhD
University of Macedonia, Thessaloniki, Greece
themis.tzimas@gmail.com

ABSTRACT: The article analyzes the role of human rights in relation to Artificial Intelligence. The main goal is to identify how human rights can contribute into a new international treaty, attempting to regulate the advances and the functions of AI, both at the present, narrow field, as well as at the level of general or super intelligence in the future. In order to do so, the article examines issues which are related to the ontology of AI, which determine the transformation of social and subsequently of legal relations too. In such a framework, the impact of human rights is presented.
KEYWORDS: Artificial Intelligence, autonomy, human rights, international law

Introduction

The present and future technological, as well as social, economic and political developments are already and will be further defined by the rise of Artificial Intelligence (Ben-Ari, Frish, Lazovski, Eldan & Greenbaum 2017, 10). Several developments – i.e. the emergence of the so- called "fourth industrial revolution" or issues related to intellectual property and patents, military operations, arts, education, medicine, governance, social policy making, finance, environment and the equivalent fields of law being some of them- indicate such a defining role.

AI explosive expansion has raised both concerns (Sofge 2015) and expectations because of historically novel and unique issues (Larson 2010, 106; Kowert 2017, 181-83).

At the core of these issues lays the unique ontology of AI, which is built on the growing and expanding autonomy of AI entities, which both complicates the relationship of AI and humans from the perspective of the latter and raises the potential for a new type of legal personhood, that of AI. In this sense, human rights become critical in terms both of a potential AI legal personhood and of humans' protection. This is the framework of the present examination.

In order to examine the role of human rights, the paper first analyses the ontology of AI. It then examines the relevance of human rights, applying them to the ontologies of AI.

1. The AI "ontology"

AI ontology is surrounded by ambiguity at a significant extent. "[I]n spite of what I regard as AI's significant achievements . . . the not so well-kept secret is that AI is internally in a paradigmatic mess" Chandrasekaran comments. (Chandrasekaran 1990, 14). The definition of AI is debatable too (Russell & Norvig 2013, 2). It has been defined as "a broad set of methods, algorithms, and technologies that make software 'smart' in a way that may seem human-like to an outside observer" (Noyes 2016) A slightly different definition describes AI as "Machines that are capable of performing tasks that, if performed by a human, would be said to require intelligence" (Scheree 2016, 363-64).

AI definitions include the elements of "consciousness, self-awareness, language use, the ability to learn, the ability to abstract, the ability to adapt, and the ability to reason" (Scheree 2016, 363-64) of goal orientation and of the rational agent (Russell & Norvig 2010, 2-3). The focus of most definitions lays in the "human- like" intelligence of machines, although that can be partially deceiving, as an entity mimicking human intelligence does not necessarily "understand" or share the patterns of human intellect (Laton 2016, 94).

AI is distinguished between weak AI, where "the computer is merely an instrument for investigating cognitive processes" and strong AI, where "[t[he processes in the computer are intellectual, self-learning processes"(Wisskirchen 2017, 10). Weak AI is labeled as Artificial Narrow Intelligence-ANI- while strong AI is further distinguished between Artificial General Intelligence –AGI– and Artificial Super Intelligence – ASI (Urban 2015). It must be noted however that ANI has already surpassed the direct control from the programmer too.

Therefore, the learning procedure and autonomy already exist having surpassed the automation phase; however until now they apply only in specific areas, unlike humans who possess general intelligence. Although AI has already"outsmarted" humans in certain, narrow areas and tasks, it cannot –yet- compete with humans, in terms of adaptable and general intelligence.

AGI will be consisted of the "type of adaptable intellect found in humans, a flexible form of intelligence capable of learning how to carry out vastly different tasks… based on its accumulated experience"(Heath 2018) enabling it to choose by itself, where and how to apply its intelligence. The "when" of AGI is debatable, although most analysts agree that within this century it will happen (Tal 2018). Super intelligence refers to the exceeding of human intelligence in the sense of"…an intellect that is much smarter than the best human brains in practically every field..."(Bostrom 1998).

While the time of the achievement of super intelligence remains at stake, its achievability is foreseen with some certainty. As an article co-authored by Stephen Hawking, Max Tegmark, Stuart Russell, and Frank Wilczek foresaw that: "…there is no physical law precluding particles from being organized in ways that perform even more advanced computations than the arrangements of particles in human brains" (Hawking et al. 2014).

The main idea is that since human brain performs computation, a different, non- biological computational entity could perform like the human brain and eventually out-perform it (Snyder- Beattie & D. Dewey 2014). At the core of AI development lays the intellectual autonomy of the entity, in combination with developments such as big data, better algorithms and improved hardware (MacDonald 2016). Intellect autonomy is built on "machine- learning", comprised of a performance and of a learning element.

The first one "senses the environment", while the latter, employs feedback from the system and amends the performance element (Marra & S. K. McNeil 2013, 1145).

Machine learning thus resembles more to "coaching" than programming (Tanz 2016; Scherer 2016, 33) and also to human learning procedure (Schuller 2017, 404). It can be also described through the cumulative contribution of three abilities: to compute information, to learn and to reason (Khoury 2017, 640).

Machine learning is already giving way–at least up to some extent–to neural networks and deep learning. Neural networks are inspired by human brain and the synapses between neurons, which function at different layers, through which, massive data run, in order to train the system. An AI neural network is a "biologically inspired computational model that is patterned after the network of neurons present in the human brain", modeling "the input-output relationship" (Nvidia 2019). Neural networks sustain and enhance machine learning, promoting and accelerating AGI.

In the framework of such procedure, AI entities need to include various components, such as logic- "as a tool of analysis, as a basis for knowledge representation, and as a programming language"(Thomason 2003) —creativity—combined with skills such as problem solving, pattern recognition, classification, learning, induction, deduction, building analogies, optimization, surviving in an environment and language processing (Hutter 2010, 125-126, 231) —communicative capacities, external knowledge, "cognitive autonomy" —in the sense of working "independently without human intervention beyond defining goals" - intuition and strategic thinking (Camett & Heinz, 2006; Suchman and J. Weber 2016, 39-40).

Machine learning and neural networks have already surpassed "rules-based programming",(Pyle & C. San Jose 2015) providing AI the capacity to function autonomously from the human programmer, surpass by far human intelligence –*currently*– in narrow, pre-determined areas, evolve and even re-programme itself. Of course, AI has not yet achieved general intelligence and is still indicating these exceptional capacities, in a "protected" environment.

Much higher autonomy will take place when AI entities will be endowed with self- awareness, in the sense of being aware of their own

existence and of placing themselves in the broader world, with –as mentioned above- adaptable intelligence which may lead to their choices not only in terms of means but also in terms of goals (Chong 2015; Schkolne 2018). Such conception of self- awareness implies a unity of subjective, mental activities, such as imaginative thinking, self- decision, creativity, self- representation and self- discovery, sentience, wakefulness, all of which tend to re- inventing one's own presence in the world. These elements describe aspects of consciousness (Herbert 1985, 249) with the latter comprehended as "…self-reflective… [as] the perception of perception, and the awareness of awareness"(Smith 1998, 281; Tegmark, 2018, 428-30, 431). Essentially, consciousness is condensed in the subjective experience, which also bears with it a certain degree of unpredictability.

Such development however should not be perceived as necessarily leading to intellect- autonomy and function, identical to that of humans. On the contrary, it is likely that the concepts of the "self" and of the surrounding environment may be inherently different for AI (Damasio 1994, 247-248). While it is with AGI and ASI that the fore- mentioned issue becomes emphatically present, it is also present with existing, AI intellect autonomy at relatively narrow fields, which can produce impressively beneficial or destructive consequences, both unpredictable and not traceable or attributable to the initial human programmer (Eden, Steinhart, Pearce & Moor 2012, 28-9; Del Prado 2015; Bostrom 2014, 26-29, 140, 155).

Summing up, the argument is that the developing ontology of AI is condensed in its expanding autonomy which tends towards subjectivity and therefore unpredictability, the extent of which is determined on the basis of intellect capacity, adaptability and generalization, as well as of autonomy. This is why the argument of the present article is that a new framework specifically designed for AI, both in its current and in its potential forms is immanently necessary.

2. A regulatory framework for AI- the role of human rights

On the basis of the above- mentioned ontological elements and of the prospect they bear to fundamentally alter human conducts or even to introduce us into an era of new "beings" and legal subjects, of non- human

orientation, the need for a legal framework, capable of present and future developments.

Until now, there are only mild and primary efforts for the establishment of a legal framework, as well as declaratory documents by private entities. Indicatively, the EU Parliament adopted a resolution about civil law rules on robotics, endorsing Asimov rules for autonomous AI and robotics (European Parliament 2017).

Other powers, such as the US, China and the UK are also working on regulatory frameworks, without having produced though coherent legal frameworks. Private institutions have contributed into the gradual formation of more de- centralized regulatory schemes, which however cannot be substitutes to full- fledged, legal schemes (Triolo P., Kania E., and Webster G., 2018; Black 2001, 103).

The answer to the question about the proper type of legal regulation must be determined on the basis of novelty, of risk and of expansion of AI. The novelty determines the extent of suitability of the existing legal systems; the risk factor, determines the prevalence of hard or soft and de- centralized law approaches; the impact, the main "beneficiaries" of the regulation; It is on the basis of a combined approach to these criteria that we reach the conclusion that novel and adaptable legal systems are required, in the sense of an international treaty so as to avoid fragmented and therefore inadequate responses (Andersen 2018, 55-56).

Existing legal systems can contribute with existing fundamental principles—albeit in some cases with the necessary changes—in order to achieve a three- end goal: preserve the safety and the rights of humans, preserve fairness among humans and when AGI will have been achieved preserve the rights which will be flowing from the potential legal personhood of AI entities. In this sense human rights, as existential rights for humans and for the international community, set the ultimate checks and balances for legal systems and therefore, potentially for the regulation of AI too (Alston 1984, 607).

Human rights can establish a regulatory framework that will be prohibiting and enabling certain AI developments and applications and also they must constitute a positive obligation of programmers, manufacturers

and owners of AI in the sense of "training" of AI systems so that they endorse the overall goals and the specific, human rights.

However, the actual implementation of human- rights' guided and trained AI will have more complexities than it seems: the growing autonomy means that the effectiveness of "training" of AI entities may eventually be proven limited and also we cannot yet foretell how a non- human, intelligent entity will comprehend in its self- development and self- conscious course, human rights. We can try and create "friendly" AI, meaning AI that will share "our" goals and our idea of humanity and of the preserve it. However we can never be absolutely certain that such guarantees will be proven efficient even in ANIand we cannot rely solely on a training procedure without a more general and intervening, regulatory framework, in different stages of AI evolution (Omohundro 2008, 483-92).

Therefore, the prospect of intelligent entities, which may be equally intelligent or superior to us, posing existential danger, could justify a slowing down or even a prohibition of certain technological advances, which lead to AGI and ASI, via a relevant treaty, establishing that AI technology that can be threatening for the superiority of human intelligence and for the goals of the international community will be prohibited (De Garis 2005, 1-2).

Such an approach however—if chosen—has the defect that it solely emphasizes upon the potential risk from AI, being therefore up to some extent, one- sided while AI applications can be double- edged; both beneficial and possibly harmful. In some sense, AI according to analysts can be proven even morally enhancing to humans (Waser 2008).

Therefore what is proposed is the intervention in advance and if needed in "correction" of the four main reasons for unethical behavior: namely "over-riding self-protection (fear); selfishness (greed); unfairness (error) on society's part; or error on the entity's part" (Waser 2008). If the ethical risk can be minimized, a general prohibition of certain AI developments will rather harm than safeguard humanity and human rights too. We need therefore to imagine a more elaborate and complicated legal system, which will be able to provide better guarantees regarding- among other areas of law- the guidance of AI by human rights as well as to capture the potentially beneficial and benevolent impact of AI, without undermining the risks too.

The first principle of such an approach must be that human rights should guide the technological research and the applications of AI, as a positive obligation of manufacturers, programmers and owners of AI to train the latter in line with human rights. Therefore, the flow of big data, the algorithms and software that are used must include human rights as part of machine learning and of the training procedure.

The second principle should refer to the differentiation among the various AI applications- actual or potential- and to technological research leading to them. It cannot be overlooked that there are applications which tend to be more beneficial for humans and for the promotion of human rights, whereas others bear more risks. Depending on the potential risk to human rights–among other things- that they represent they can be divided between low, medium and high risk AI.

Such categorization can be determined on the basis of the goals, as well as of the means and will be leading to policies of further promotion, of partial restriction or of prohibition of certain applications—actual or future—and of technology leading to them, depending on the risk that they pose. There may be several and different policies and measures, such as the control of the type of data provided or the disconnection of certain AI applications from the cyberspace or parts of it.

The third principle of a potential legal regulation, on the basis of human rights, engulfs the most intriguing issue, which is that of the regulation of the potential emergence of AGI and of ASI. Can the path towards such developments be legitimate under human rights imperatives? The answer to the question is pre- legal: if the prevalent assumption is that AGI or and ASI will certainly or likely become hostile towards humans, then human rights impose the obligation to terminate research moving towards this direction, at least "one step" before reaching any of these two levels. Otherwise, we must focus upon these checks and balances, in accordance with human rights so that we keep it non- hostile and beneficial for us, enhancing its benevolent tendency.

In case however AGI and ASI is eventually achieved, human rights will have to adapt given that most likely there will be an international or- to better present it- a global community comprised from human and non-

human being of equal or superior intellect capacity. While human rights may be able to retain their relevance for humans they will stop constituting the fundamental norms of that new, global community.

One last thing that remains to be discussed is how human rights will be related with the potential legal subjectivity of AI, in case the latter is achieved (Lawson 1957, 915; Solum 1992, 1285; Barrat 2013, 39-41; Dowell 2018, 321, 327-29). In this sense, all legal systems are human- centric and take for granted that humans are the dominant and more developed form of being- intellectually speaking- the welfare of who constitutes the main goal. The impetuous development of AI can challenge this, until now, self- obvious fact (Anderson M. & Anderson S. L. 2011, 7-13).

Up to the extent that conscience, reason, self- awareness and intellect autonomy will be identified with non- human beings as well, aspects of or a complete legal personhood may be attributed to them too (Bayern 2015, 104). What complicates things is that defining factors of human personhood which fundamentally shape legal subjectivity and therefore legal systems too- for example death or the way we comprehend life, physical harm and danger, relative equality, relative cultural homogeneity among humans - may be irrelevant or at least will be adjusted seriously, when applied in AI entities (Khoury 2017, 646).

The lack of fear of sanction and the ability to replicate them, imply foundations and existential ideas which are completely different from the ones upon which legal systems until now are built (Scherer, 2016, 367). In other words, we cannot foretell how subjectivity and its legal aspect will be experienced by AGI and ASI and therefore their potential legal behavior of AGI and ASI remains as we speak at large terra incognita. What in principle can be foreseen is that legal personhood will be analogical to growing autonomy. AI entities will have an evolving, most likely at some stages a partial or limited and sui generis type of legal personhood (Watson 2018, 68), which may develop through AGI and ASI into a complete one.

On the basis of such assumptions we can foretell "two-plus-one" potential layers of legal personhood: the one emerges out of the self-awareness or the existential awareness of AI entities; the second emerges out of the interactions of AI entities with existing legal persons, referring to

the vast area of AI applications and attempting to safeguard existing legal persons' rights, the relationships among them and the rights of AI entities; the additional layer refers to the interaction of AI entities with political communities or to the formation of "political communities" by AI entities themselves, on the basis of the potential for self- organization of fully autonomous AI entities (Ahmed & Glasgow 2012).

The first layer can be formulated by rights flowing out of the self-preservation of entities which possess self- awareness and consciousness. Not only for terminological but also for substantial reasons we cannot speak about human rights of AI entities. Nevertheless it is interesting to notice the UDHR guarantees human rights on the basis not only of the common interest to preserve peace but also – in existential terms – of the endowment of humans with reason and conscience.

Rights related to existence, conscience, self- preservation, to autonomy-liberty and freedom- and self- enhancement can be relevant with and suitable for fully autonomous, AI "beings", which will have reached the level of AGI or/and ASI. A set of existential rights may gradually develop in the sense of fundamental AGI and ASI rights, including the preservation of existence, intellectual development and to rights flowing out of AGI and ASI creations and activity.

The second layer is consisted of the need to design a legal system capable of preserving fairness, social and political rights and therefore human rights, among humans in light of the different uses and applications of AI, as well as on the basis of AI unique legal subjectivity; in this sense it should also be able to preserve fairness for AI too though.

The issue is condensed at large in matters of liability, ownership, and of profitability because of AI creations. The complexities arise because of the growing autonomy of AI which means that it is not always easy or even possible to trace the human control behind AI entities' creations, both when liability and responsibility must be determined as well as when profit is to be shared (Childers 2008, 128).

Liability and ownership touch upon the issues of reparation and restitution, whereas of profitability on the issues labor, social and indirectly

political rights and therefore they are linked with human rights' goals –such as fairness and dignity- as well as with specific rights.

The former refer to the need to identify responsibility over AI entities' actions and omissions. An initial approach can be to hold the owner or the programmer of the autonomous AI system liable for the latter's potential wrongful conducts. Such a solution may seemingly provide some extent of legal certainty, in the sense that the owner has knowingly accepted the potential dangers from the unpredictability of the entity. However, relying solely on such ground, when referring of course to fully autonomous AI entities, eventually could bear the seed of unfairness, due to the level of unpredictability and self- development of the AI entity (Moravec 2009).

The counter -arguments suggest that the above approach fails to capture the essence of deep- learning procedures and of how the latter overcomes the initial programming, (Grimmelmann 2016, 408) establishing both creativity and autonomous intellect for AI entities, even in relation to ANI and far more with AGI and ASI. Therefore, the recognition of AI entities as autonomous creators is proposed (McFarland, 2016).

On the basis of this latter perception, a different approach is to transfer the burden of responsibility to the AI entity itself. From this perspective it is through AI entities that restitution must come because their autonomy exceeds automation and human control.

Such a legal regulation could entail "corrective" measures on an AI entity or reparation from AI entities through their creations. Matters of restitution will profoundly emerge. A solution can be a public or/and private insurance scheme, established with a compensatory rationale- i.e. in exchange for the public access to autonomous AI entities' creations (McLean 2002, 205). The most suitable approach may be a combination of aspects of the two, above- mentioned proposals, depending on the level of autonomy; a multi-level approach, which will entail—cumulatively or alternatively—and on the basis of the level of autonomy of the entity, liability of the manufacturer or of the programmer—in "hardware cases" and in "software cases" respectively—when the autonomy of the entity is lower and the human programmer, manufacturer or owner may be more directly or indirectly "traceable".

The extension of autonomy shifts gradually the burden of responsibility to AI itself. In this framework, a scheme of restitution out of AI creations, a public / private insurance scheme and corrective measures in the algorithms, software and training of AI entities can be imposed.

Similar issues arise in relation to profitability out of the legal status of autonomous AI entities' creations. The question is if it is humans or the autonomous AI entities themselves that should profit out of the latters' creations or whether some other legal framework should be adopted.

One approach is that the ownership and profits from AI entities creations must be attributed to humans- the initial programmer, the owner or the user of the entity. It invokes in its favor, the unfamiliarity of AI entities with profit as well as their supposed ellipsis of the necessary "creative spark" or of "inventive concept" (Abott 2016, 1079- 1082, 1086-1099) in order for the latter either to be provided profit or to be recognized as autonomous creators. Parenthetically such arguments invoke that profitability is related to IP rights theories, which are essentially human- centered (Pearlman 2018, 20-35).

The opposite arguments suggest that the above approach fails to capture the essence of deep-learning procedures and of how the latter overcomes the initial programming, establishing both creativity and autonomous intellect for AI entities. As we already know, AI applications such as Alpha Go or arts' applications already demonstrate some extent of creativity. This characteristic will be further developed in AGI and ASI. We may not be able to foretell and determine the nature of AI creativity or of manifestations of creativity in its future development but that does not stop us from understanding that there are certain AI acts which do not constitute the outcome of human act and which are not controlled by humans. After all, not even human creativity is completely "de- codified". Therefore, AI entities can be recognized as autonomous creators being attributed a subsequent legal subjectivity generating the equivalent rights.

Such AI rights could be considered as "inspired" by social rights. The ontological identities of AI however make it difficult to draw analogies, regarding social rights between humans and AI because we cannot foresee if there will be any type of "social" organization of AI as well as what that may be.

It seems at this point however unlikely those AI entities will be in need of some type of wealth accumulation. Therefore, it can be argued that fairness among humans, AI ontology and legal subjectivity justify not a -fundamentally irrelevant with AI entities' -attribution of ownership or IP rights to AI, but due to the recognition of AI as creators, the placement of such creations (Bakry & He, 2015), in the framework of the public space, as freely accessible, maximizing their social utility (Litman 1990: 968-1022).

Again, international law can contribute into the formation of a legal framework, serving the fore- mentioned objectives from the perspective both of human and of AI legal subjectivity, on the basis of fairness, which lies at the foundations of human rights, as well as on the basis of numerous other specific, human rights (Tsagourias 2015, 25).

The important remark is that it is difficult to authoritatively comprehend the legal subjectivity of AI, especially as AI autonomy evolves. Human rights as a concept cannot be applied to AGI and ASI legal subjectivity. They may be used however as a guide in the uncharted waters within which a new legal system will have to sail if AGI and ASI become reality.

Conclusions

The present article addressed AI and cyberspace initially from their ontological perspective in order then to assess how the latter influence the current and the potential, future legal debate. The fundamental elements of AI ontology are its evolving autonomy and intellect capacity and the potential of these characteristics to reach an intellect level, equal or even superior to human, whereas of cyberspace are its ecumenical expansion, the merging of physical and cyber world and the movement in its framework with the speed of electron. Both of them present unique challenges to existing legal systems already. Their development and their merging however bears the potential of a completely novel landscape at all levels of human conduct and therefore at the legal level too.

The argument of the article is that on the basis of different criteria, a new international treaty is needed which will be based at large on human rights and will be able to establish or at least start constructing a type of international rule of law for both AI and the cyberspace.

The main goals must be to preserve human rights for humans and fairness among them, in light of AI and cyberspace applications but also to provide us, on the basis of some analogies, an insight about how rule of law should be adjusted on the basis of new, emerging legal subjectivity of AI.

Human rights must play the role of the fundamental pillar of an effective legal system which will promote or discourage certain AI technological research and applications, on the basis of the danger that they pose for human rights, not submitting to pessimistic views about AI but without underestimating the dangers either.

In addition, the international community, when presented with the dilemma of legitimizing or not the emergence of AGI and ASI will have to take into account whether the latter can be "controlled" in the sense of not endangering human rights or not.

Eventually however, what cannot be done for human rights is to be absolutely safeguarded in a potential, future situation of equally intelligent entities and therefore legal subjectivities, or in a situation within which humans will not be the superior entities intellectually. Such entities could lead to a moment of legal singularity- in analogy with the moment of singularity for AI in general, when new legal systems, with new types of rights will be needed. Even before that "moment" however, the issue of legal subjectivity of AI even in narrow areas will emerge. It is in such a framework that human rights can lead us to a rule of law at least until- and if- AGI and ASI emerge.

References

Abbott, R. 2016. "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law". *B.C.L. Rev*. 57: 1079-126.

Ahmed, H. & Glasgow J. 2012. "Swarm Intelligence: Concepts, Models and Applications: Technical Report," 2012-585, *Queen's Univ. School Of Computing*, 2 February, ftp.qucis.queensu.ca/TechReports/Reports/2012-585.pdf, archived.

Alston, P. 1984. "Conjuring up new human rights: a proposal for quality control." *American Journal of International Law* 78: 607-21.

Andersen, L. 2018. "Human Rights in the Age of Artificial Intelligence." *Accessnow org*, www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf.

Anderson, M. & Anderson S. L. 2011. *Machine Ethics*. Cambridge: Cambridge University Press.

Barrat, J. 2013. *Our Final Invention: Artificial Intelligence and the End of the Human Era*. New York: St. Martin's Press.

Bayern, S. 2015. "The Implications of Modern Business Entity Law for the Regulation of Autonomous Systems." *Stan. Tech. L. Rev*. 19: 93-112.

Ben- Ari, D., Frish Y., Lazovski A., Eldan U. & Greenbaum D. 2017. "Danger, will Robinson? artificial intelligence in the practice of law: an analysis and proof of concept experiment." *Richmond Journal of Law and Technology* 23: 3- 53.

Black, J. 2001. "Decentring Regulation: Understanding the Role of Regulation and Self-Regulation in a 'Post-Regulatory' World." *Current Legal Probs*. 54:103-46.

Bostrom, N. 1998. How Long Before Superintelligence? Oxford Future of Humanity Institute, Faculty of Philosophy & Oxford Martin School, University of Oxford, https://nickbostrom.com/superintelligence.html.

Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Camett, J. B. & Heinz E. 2006. "John Koza Built an Invention Machine." *Popular Science*, Apr 19, www.popsci.com/scitech/article/2006-04/john-koza-has-built-invention-machine.

Chandrasekaran, B. 1990. "What Kind of Information Processing is Intelligence?" In *The Foundations of Artificial Intelligence*, edited by D. Partridge & Y. Wilks. Cambridge: Cambridge University Press.

Childers, S. J. 2008. "Don't Stop the Music: No Strict Products Liability for Embedded Software." *U. Fla. J.L. & Pub. Pol'y* 19: 125-49.

Chong, C. 2015. "This robot passed a 'self-awareness' test that only humans could handle until now." *Tech Insider*, July 23, www.businessinsider.com/this-robot-passed-a-selfawareness-test-that-only-humans-could-handle-until-now-2015-7.

Damasio, A. R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.

De Garis, H. 2005. *The Artilect War: Cosmists Vs. Terrans*. Palm Springs, CA: ETC Publications.

Del Prado, G. M. 2015. "Stephen Hawking Warns of an 'Intelligence Explosion.'" *Business Insider*, October 9, www.businessinsider.com/stephen-hawking-prediction-reddit-ama-intelligent-machines-2015-10, archived at https:// perma.cc/P4NL-2AJ2.

Dowell, R. 2018. "Fundamental Protections for Non-Biological Intelligences or: How we Learn to Stop Worrying and Love our Robot Brethren." *Minnesota Journal of Law, Science & Technology* 19: 305-36.

Eden, A., Steinhart E., Pearce D. & Moor J. 2012. Chapter I, Singularity Hypotheses: An Overview, in *Singularity Hypotheses, A Scientific and Philosophical Assessment*, edited by Eden, J. Moor, J. Soraker & E. Steinhart, New York Springer.

European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

Faried, Bakry M. & He Z. 2015. "Autonomous Creation - Creation by Robots: Who owns the IP Rights?" Mar. 15, *Maastricht University Blog Intell. Prop. & Knowledge Mgmt*. law.maastrichtuniversity.nl/ipkm/autonomouscreation-creation-by-robots-who-owns-the-ip-rights/.

Grimmelmann, J. 2016. "There's No Such Thing as a Computer-Authored Work - and It's a Good Thing, Too." *Colum. J. L. & Arts* 39: 403-16.

Hawking, S., Tegmark M., Russell S., and Wilczek F. 2014. "Transcending Complacency on Superintelligent Machines." *Huffpost*, https://www.huffingtonpost.com/stephen-hawking/artificial-intelligence_b_5174265.html?ec_carp=3359844804041712164.

Heath, N. 2018. "What is AI? Everything you need to know about Artificial Intelligence." *ZDNet*, February 12, https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/.

Herbert, N. 1985. *Quantum Reality: Beyond the New Physics*. New York: Anchor Books Editions.

Hutter, M. 2010. *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*, Springer.

Khoury, A. 2017. "Intellectual property rights for 'Hubots': on the legal implications of human-like robots as innovators and creators." *Cardozo Arts and Entertainment Law Review* 35: 635-69.

Kowert, W. 2017. "The foreseeability of human-artificial intelligence interactions." *Texas Law Review* 96, 181, at pp. 181-204.

Larson, D. A. 2010. "Artificial Intelligence: Robots, Avatars, and the Demise of the Human Mediator." *Ohio State Journal on Dispute Resolution* 25: 105-64.

Laton, D. 2016. "Manhattan_Project.Exe: A Nuclear Option for the Digital Age." *Catholic University Journal of Law & Technology* 25: 94-153.

Lawson, F.H. 1957. "The Creative Use of Legal Concepts." *N.Y.U. L. Rev.* 32:909.

Litman, J. D. 1990. "The Public Domain." *EMORY Law Journal* 39: 965- 1023.

MacDonald, F. 2016. "Harvard Scientists Think They've Pinpointed the Physical Source of Consciousness." *Sci. Alert,* Nov. 8, www.sciencealert.com/harvard-scientists-think-they-ve-pinpointed-the-neural-source-of-consciousness.

McFarland M. 2016. "Google's Computers Are Creating Songs. Making Music May Never Be the Same." *Washington Post*, June 6, www.washingtonpost.com/news/innovations/wp/2016/06/06/googles-computers-are-creating-songs-makingmusic-may-never-be-the-same/?utm_term=.55226125405c.

McLean, T. R. 2002. "Cybersurgery-An Argument for Enterprise Liability." *J. Legal Med*. 23: 167-210.

Marra, W. C. & McNeil S. K. 2013. "Understanding 'The Loop': Regulating the Next Generation of War Machine." *Harvard Journal Law. & Public. Policy* 36: 1139-87.

Moravec, H. 2009. "Rise of the Robots: The Future of Artificial Intelligence." *Scientific America*, Mar. 23, www.scientificamerican.com/article.cfm?id=rise-of-the-robots.

Noyes, K. 2016. "5 things you need to know about A.I.: Cognitive, neural and deep, oh my!" *Computerworld*, March 3,www.computerworld.com/article/3040563/enterprise-applications/5-things-you-need-to-know-about-ai-cognitive-neural-anddeep-oh-my.html [http://perma.cc/7PW9-P42G.

Nvidia, Artificial Neural Networks. 2019. Available at https://developer.nvidia.com/discover/artificial-neural-network.

Omohundro, S. 2008. "The Basic AI Drives, in Artificial General Intelligence." In *Proceedings of the First AGI Conference*, edited by P. Wang, B. Goertzel & S. Franklin, Amsterdam, IOS.

Pearlman, R. 2018. "Recognizing Artificial Intelligence (AI) As Authors And Inventors Under U.S. Intellectual Property Law." *Richmond Journal of Law and Technology* 24: 2-38.

Pyle, D. & San Jose C. 2015. An executive's guide to machine learning, McKinsley Quarterly, June, https://www.mckinsey.com/industries/high-tech/our-insights/an-executives-guide-to-machine-learning.

Russell, S. J. & Norvig P. 2013. *Artificial Intelligence: A Modern Approach*. New Jersey: Pearson Education Limited.

Scherer, M. U. 2016. "Regulating Artificial Intelligent Systems: Risks, Challenges, Competences, and Strategies." *Harvard Journal of Law & Technology* 29: 353-99.

Schkolne, S. 2018. "Machines Demonstrate Self-Awareness." *Medium*, April 4, https://becominghuman.ai/machines-demonstrate-self-awareness-8bd08ceb1694.

Schuller, A. 2017. "At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law." *Harvard National Security Journal* 8: 379-424.

Smith, JC. 1998. "Machine Intelligence And Legal Reasoning." *Chicago-Kent Law Review* 73:277.

Snyder - Beattie, A. & Dewey D. 2014. "Explainer: what is superintelligence?" *The conversation*, July, 18, https://theconversation.com/explainer-what-is-superintelligence-29175.

Sofge, E., 2015, "Bill Gates Fears A.I., But A.I. Researchers Know Better", *Popular Science* January 30, www.popsci.com/ bill-gates-fears-ai-ai-researchers-know-better.

Solum, L. B., (1992), Legal Personhood for Artificial Intelligences, N.C. L. REV., 70: 1231-89.

Suchman, L. and Weber J. 2016. "Human-Machine Autonomies." In *Autonomous Weapon Systems: Law, Ethics, Policy*, edited by N. Bhuta, S. Beck, R. Geib, H. Yan Liu and C. Kreb. Cambridge: Cambridge University Press.

Tal, D. 2018. "Forecast | How the first Artificial General Intelligence will change society: Future of artificial intelligence P2." *Quantumrun special series*, May 24, https://www.quantumrun.com/prediction/first-artificial-general-intelligence-society-future.

J. Tanz. 2016. "Soon We Won't Program Computers. We'll Train Them Like Dogs." *Wired*, May 17, www.wired.com/2016/05/the-end-of-code/.

Tegmark, M. 2018. *Life 3.0, Being Human in the Age of Artificial Intelligence*. London: Penguin Books.

Thomason, R. 2003. "Logic and Artificial Intelligence." *Stanford Encyclopedia of Philosophy*, plato.stanford.edu/entries/logic-ai/, archived at https://perma.cc/3RPH-PVKV.

Triolo, P., Kania E., and Webster G. 2018. "Translation: Chinese government outlines AI ambitions through 2020." *New America*, January 26

https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-government-outlines-ai-ambitions-through-2020/.

Tsagourias, N. 2015. "The legal status of cyberspace." In *Research Handbook, International Law and Cyberspace*, edited by N. Tsagourias & R. Buchan. Londonn: Edward Elgar Publishing.

Urban, T. 2015. "The AI Revolution: The Road to Superintelligence." *Wait But Why*, Janauary 22, waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html.

Waser, M. 2008. "Discovering The Foundations Of A Universal System of Ethics As A Road To Safe Artificial Intelligence." *AAAI*, www.aaai.org/Papers/Symposia/Fall/2008/FS-08-04/FS08-04-049.pdf.

Watson, B. 2018. "A Mind of its Own--Direct Infringement by Users of Artificial intelligence Systems." *IDEA: J. Franklin Pierce for Intellectual Property* 58: 65-93.

Wisskirchen, G. et al. 2017. "Artificial Intelligence and Robotics and their Impact on the Workplace" April 2017, p. 10. London: IBA Global Employment Institute.