

Introduction to the special issue on machine law

Bartosz Brożek¹ · Jaap Hage² · Bipin Indurkhya³

Published online: 7 September 2017

© The Author(s) 2017. This article is an open access publication

We are surrounded by machines. From simple ones—AC motors and transformers—through radio receivers, TV sets, smartphones and personal computers, to sophisticated AI systems, such as self-driving cars, autonomous weapons and IBM’s Watson. The advances in technology have reshaped the world we inhabit, including our social environment. When iPhone is the girl’s best friend, our communication and decision-making is aided by complex algorithms, and various tasks so far reserved for human beings are carried out by robots, the contemporary societies are not what they used to be. Moreover, the technology is advancing at such a rapid pace that many ideas, such as companion and sex robots, which used to be a fodder for science fiction are fast becoming a reality.

This is a profound challenge for any legal system. The law is there to regulate the actions of individuals so that they contribute to the functioning of large societies. It means that legal institutions should be designed in such a way as to embrace any changes and developments that reshape our communal practices. For this reason, technological progress has been a focus of lawyers’ debates since the first industrial revolution. The great discoveries of the nineteenth and twentieth centuries—car, airplane, radio, TV, computer, the Internet—have not only influenced the existing legal institutions, but have also led to the establishment of entirely new branches of law. Arguably, however, they did not revamp the very foundations of their contemporary legal systems, but served as a means for regulating interactions between human beings. Technology has been considered only as a tool used by

✉ Bipin Indurkhya
bipin@agh.edu.pl

¹ Faculty of Law and Administration, Jagiellonian University, Cracow, Poland

² Faculty of Law, Maastricht University, Maastricht, The Netherlands

³ Institute of Philosophy, Jagiellonian University, Cracow, Poland

human actors—a tool capable of changing the nature of our interactions, but a tool nevertheless.

This situation has changed dramatically with the introduction of autonomous machines, which are reactive (they respond in a timely fashion to changes in the environment), autonomous (they exercise control over their own actions and are not directly controlled by any other agent), goal-oriented (they act in a purposeful way and do not simply react in response to the environment), and temporally continuous (they are always running). The question emerges whether—from the legal perspective—such machines should remain ‘tools’ in the hands of human actors, or whether they should rather be considered genuine legal patients or agents. This problem lies at the very heart of the law: should we start thinking about reconceptualising the foundations of our legal systems, granting autonomous machines the status (or, at least, a partial status) so far reserved for human beings?

The papers collected in this special issue of ‘Artificial Intelligence and Law’ all address some aspect of the aforementioned problem. Three of them—J. Hage’s, J. J. Bryson, TD Grant, and ME Diamantis’, and B. Brożek and M. Jakubiec’s—attempt to spell out the conditions for granting autonomous machines the status of a legal agent. Hage argues that it is possible to hold autonomous agents themselves, and not only their makers, users or owners, responsible for their acts. He claims that there are no metaphysical or conceptual barriers to make such an attribution of agency impossible, and that the question of whether autonomous systems should be considered legally responsible is a purely utilitarian question: if such a legal manoeuvre is considered beneficial, it would be fully justified.

On the other hand, Bryson, Grant and Diamantis argue that it is not contestable that autonomous machines can be granted legal personhood, since it is a conventional conceptual construct, and that the decision to do so should be determined purely by its consequences. However, they further argue that the potential costs of granting autonomous systems the status of a legal agent seem to outweigh the foreseeable benefits.

Brożek and Jakubiec take a slightly different stance. They also acknowledge that it is technically possible to consider autonomous machines as legal agents; however, they claim that such a manoeuvre would be ineffective for conceptual reasons. The conceptual apparatus regarding legal responsibility is well rooted in folk psychology (the way people conceptualise, understand and explain their actions and the actions of other people), and it is difficult to see how the actions of artificial agents can be incorporated into the folk-psychological model of agency.

In their paper, L. Frank and S. Nyholm consider a more concrete problem connected to the agency of autonomous machines: whether it is conceivable, possible, and desirable that humanoid robots should be designed in such a way that they are capable of consenting to sex? They discuss reasons for both positive and negative answers to this question, taking into account such problems as the concept of consent in general, and the relationships between consent and free will and between consent and consciousness.

The following three papers deal with a different aspect of the main problem addressed in this volume—what should be the inner architecture of autonomous machines so that they may follow the law and be considered legally (or morally)

responsible for their actions. F. Podschwadek considers the question of what would be the requirements of an autonomous moral agent. He argues that a full moral autonomy implies the option of deliberately acting immorally, not merely through an error in identifying the morally correct action in a given situation. In other words, such artificial moral agents would have the potential for moral fallibility, i.e. for rejecting the moral system they are to follow altogether.

H. Prakken, in turn, considers the main problems connected with designing autonomous vehicles that respect the traffic law. He observes that traffic regulations, although quite simple and precise in comparison to other areas of the law, generate a number of troublesome issues for an artificial system. They include vagueness, specific and general exceptions, and the role of the principles of civil liability as providing indirect cues for the behaviour of an autonomous vehicle. Further, Prakken describes three approaches to developing the logical architecture of an autonomous vehicle (regimentation, reasoning, and learning), and discusses the abilities an autonomous vehicle must have in light of the requirements of the traffic regulations (e.g., complex object recognition). He also addresses the problem of knowledge representation, and highlights the difficulty connected with interpreting legal provisions.

Finally, G. Contissa, F. Lagioia, and G. Sartor address the problem of legal responsibility connected to accidents involving autonomous vehicles. In particular, they consider some scenarios in which an autonomous vehicle faces a situation similar to the notorious trolley problem. They claim that such a situation would lead to serious troubles when it comes to ascribing legal responsibility, and propose to remedy them by equipping the autonomous vehicle with a device (The Ethical Knob), which would enable its user to choose the ‘ethical mode’ of the car’s behaviour (e.g., egoistic, impartial or altruistic). In this case, the ‘decisions’ of the vehicle would—ultimately—be the decisions of the user, making the ascription of criminal liability possible. They also consider a more complex solution when the Ethical Knob has a continuous rather than a discrete setting.

We hope that the papers collected in this volume will contribute to the ongoing debates pertaining to the legal status of autonomous machines. We thank all the contributors and the reviewers for their effort and cooperation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.