

Robots, AI, and the question of 'e-persons' — a panel at the 2017 Science in Public conference, 10–12 July 2017

Michael Szollosy

Abstract

In response to EU draft legislation on robots and artificial intelligence — which included the headline-grabbing proposals to introduce rights for 'e-persons' and necessitating that robots come equipped with a 'kill switch' — a diverse group of experts and academics gathered in Sheffield as part of the Science in Public 2017 conference. Panellists and the audience discussed the origins and implications of the ideas behind the EU initiative, and more specifically, whether robots or artificial intelligence qualifies for right as 'persons', and how the EU proposal imagines robots and artificial intelligence in particular, historically-contingent ways that influence or distort our present discussions and attempts to legislate on the future use and development of technology.

Keywords

Public perception of science and technology; Public understanding of science and technology; Science and policy-making

In January 2017, the European Union Legal Affairs committee voted to adopt (by 17 to 2, with 2 abstentions) the resolutions in a report calling for the European Commission to put forward new rules for the legal governance of robots and artificial intelligence.¹ Though the report contained some very wide-ranging and, some say, radical proposals, the vote itself was merely one stage in a long and on-going process to re-evaluate a set of laws and guidelines with regard to robots and artificial intelligence. The British newspapers, however, particularly the tabloid press (notorious for bad science reporting and hostile to most things coming out of the EU) seized upon this announcement and offered their readership some dramatic headlines. *The Daily Mail*, always reliably panic-stricken, though as Britain's second-highest selling newspaper (and the most visited news website in the world²) not something that can be ignored in terms of shaping public opinion, greeted the news with the headline,

¹<http://www.europarl.europa.eu/news/en/news-room/20170110IPR57613/robots-legal-affairs-committee-calls-for-eu-wide-rules>.

²<http://www.bbc.co.uk/news/magazine-16746785>.

'Robots should be given legal status as "electronic persons" and must be fitted with "kill switches" to prevent a Terminator-style rise of the machines, warn EU MEPs'³

The article featured not one but *three* pictures of the genocidal, bipedal Terminator robot, despite the fact that this is entirely fictional robot and not at all mentioned in the Legal Affairs Committee's report.

In March, Tony Prescott, Director of Sheffield Robotics, and I, a Research Fellow on the social and cultural impacts of robotics and AI, issued a call for papers in order to arrange a panel at the Science in Public 2017 conference that was to be held in Sheffield in July. The theme for this year's conference was 'Science, Technology and Humanity' and, specifically, how rapid scientific and technological change force us to question what it means to be human.

We proposed a panel that would look in more detail at this category of 'electronic persons' mentioned in the EU report: the feasibility, the usefulness (or otherwise) of the idea and the implications (social, economic, ethical, philosophical) for both these new electronic persons and the more traditional, fleshy sort. We sought to understand the concept of 'electronic personhood', in its specific (and potential future) contexts in legislation, and in the context of the report's wider recommendations and for human societies more generally.

We received many proposals for papers, and organised the papers to present two sessions at the Science in Public conference on 10 July, 2017. I introduced proceedings and delivered the first paper, beginning with a summary of the Legal Affairs Committee's report and some examples of reactions to the report in the popular media. I explained how, despite being potentially valuable in clarifying some of the legal issues surrounding robots, the entire report and, more specifically, the notion of e-persons, are problematic in terms of their contexts: in situating robots and AI within a particular historical and ideological space, from Frankenstein to Asimov, the report reproduces very old and largely inaccurate conceptualisations of both robots and human beings. I exposed these historical origins, going back to the Enlightenment and how ideas of 'automatons' (and later, fictional robots) came to be touchstones through which we tried to understand our changing relationship with technology. The EU report therefore fails to regard robots and AI in a useful, accurate way and instead uses robots as a proxy for more fundamental (and largely unspoken) debates about what it means to be human.

I explained how this approach to policy-making on robotics, however well-intentioned, exacerbates public fears about robots, and enables this long-familiar narrative of robots as the nemesis of humanity. This popular conception demonstrates how robots become containers for cultural anxieties about what it means to be human, and how such anxieties have a negative impact particularly on the potential beneficial impacts of social robotics, robots as they could be employed in education or care. I concluded with an assessment, in this light, of the recommendations that are before the EU Commission, and make some concrete recommendations of my own as to how they might better be re-imagined for the benefit of robots and humanity.

³<http://www.dailymail.co.uk/news/article-4113674/Robots-fitted-kill-switches-prevent-Terminator-style-rise-machines-warn-EU-MEPs.html#ixzz4grk8e2Sv>.

Next to present was Aida Ponce Del Castillo from the European Trade Union Institute in Brussels. As a lawyer with an in-depth knowledge of the structures and operations of the European Parliament, she began by clarifying the status of the report, and how it might progress through the European bureaucracy, for which all the panel participants were all grateful. (For the record, the EU Parliament passes a resolution, which is then presented to the EU Commission. Only the Commission has executive power to make the Parliament's resolution law. The Commission can decide to do something, to do nothing or simply standstill. They have three months to reply, and since that time has already passed it seems as though they are deciding to standstill, so it looks unlikely this report will be made into European law. It seems that the Commission are much more concerned right now with data protection and machine safety.) So, Ponce Del Castillo does not think that the resolution will lead to new legislation across the EU, but it has provoked very important debate (as our panel demonstrates).

Ponce Del Castillo was concerned that even after this report we still don't know what a 'robot' is; the EU draft legislation does not offer a clear definition. While such a definition is badly needed, she also warned at the same time against too rigid a definition, as constant changes in the technology means that there is a risk of creating obsolete categories and regulatory traps. (She cited the case of EU legislation on nanotechnology, where there are now static definitions that few are happy with and do not work with emerging technology.)

Ponce Del Castillo explained on the issue of electronic personhood that if robots and artificial intelligence are to be considered legal persons they would acquire rights *and* obligations, necessarily becoming responsible and accountable. The important questions and tricky detail would then be what rights and obligations would be. She also warned that in giving robots liability, as such a move takes liability away those from behind the technologies, potentially exonerating designers, engineers and corporations from responsibility for their creations. (This question was raised often and urgently in the discussion, so it is clearly something that is of general concern.)

In contrast to such regulation, Ponce Del Castillo looked at the potential application of 'soft measures' in the control of robots and artificial intelligence. However, while codes of conduct can be useful, they are not instruments of governance. Certification and technical standards are other possibilities, but both operate on the basis of members/producers setting their own standards, and exclude participation from those outside a narrow, closed system. Moving forward, then, Ponce Del Castillo recommends greater visibility, a registry to ensure transparency, and a collaboration between all actors and users.

The next to speak, and the final paper in the first session, was Robert Gaizauskas from the Department of Computer Science at the University of Sheffield, who tried to answer the question 'Can robots be e-persons?' in collaboration with William Sweet, professor of philosophy from St Francis Xavier University in Nova Scotia, Canada. Gaizauskas began his talk with the conclusion that yes, robots — or, as he preferred, DIAs, digital intelligent agents — can be e-persons, and therefore are entitled to some rights. Gaizauskas cited the case of hitchBOT, the robot designed by Canadian researchers that 'hitchhiked' across Canada in 2014 (and parts of Germany in 2015), but was destroyed when trying to hitchhike across the United

States.⁴ Without setting aside the emotional and cultural issues that hitchBOT gave rise to, Gaizauskas asked whether hitchBOT's rights were damaged, or whether this is simply a case of property damage.

Gaizauskas pointed out that the category of 'human' and 'persons' aren't actually identical (as evident, for example, in the case of corporations, or humans that are kept alive entirely by artificial means). Gaizauskas also demonstrated that we ascribe different kinds of rights to different kinds of beings, so he took a more basic question as his starting place: why does a being have rights? Although space does not permit here a full examination of all 6 models that Gaizauskas offered in answer to this question, it is clear that robots can be granted rights based on several of the constituent criteria. To summarise some of the main points, robots, or DIAs, can have rights

- because they share many properties and characteristics with other non-human (or non-living) entities that are already granted rights
- because they have interests independent of their makers
- because in executing their programs, DIAs function in ways analogous to many non-conscious entities (e.g. plants)
- because they can initiate certain processes, and therefore can 'act', so should be considered 'agents', at least to some degree
- because they can evoke sympathetic reactions to their conditions (e.g. hitchBOT)

Furthermore, robots/DIAs can have obligations because they can be considered to have *interests*, are able to identify alternative courses of action, identify their own interests and those of others, reflect upon the likely impact of its actions, carry out intended actions and adapt their reasoning and behaviour as new information becomes available. On the question of what sort of obligations robots/DIAs might have, Gaizauskas listed an obligation to respect human rights and human life (as corporations must do), and an obligation to help when there is no risk to the robot's survival, but *not* an obligation to follow the law, because sometimes one has a moral obligation not to follow the law.

The second session began with a presentation from Jonathan Penn, a doctoral candidate from Cambridge University in the History and Philosophy of Science. His historical contextualisation centred on the birth of artificial intelligence and the 1956 Dartmouth Summer Research Project (where the term 'AI' was coined). Penn's particularly centred on Herbert Simon, the political scientist, economist and sociologist. Penn demonstrated how Simon's idea of *bounded rationality* and his work at the Rand Corporation, where (with Allen Newell) he used computers to model human decision-making, shaped the early conceptualisations of artificial intelligence.

Because of the key role these ideas played in the foundations of the ideas of artificial intelligence, and how this history still implicitly informs our present

⁴<https://www.theguardian.com/technology/2015/aug/03/hitchbot-hitchhiking-robot-destroyed-philadelphia>.

conceptualisations, Penn asserts that artificial intelligence — and the question of e-persons — cannot be separated from the question of public administration. As with my talk, Penn demonstrated how a (largely hidden) history of artificial intelligence is still shaping our conversations and attempts at legislation. In the case of e-persons, as with other attempts to create an ethics and rules for the governance of robotics and AI, what is also clear is the extent to which these ideological histories still play a vital role in how we use and perceive technologies; as a particularly intriguing example, Penn cited Blackrock (the world's largest asset manager) and its quest to employ artificial intelligence in financial markets.

The final talk of the session was delivered by Tony Prescott, from the Department of Psychology at the University of Sheffield. Prescott began with the question, what is a person? Turning to Locke and Daniel Dennett for answers, Prescott proposed that a person is a being with reason and language, a being capable of possessing mental states such as beliefs, capable of relationships and morally responsible for actions, and someone who is treated as a person by others. None of these qualities, Prescott points out, requires an actual, material body. Prescott's answer to this question shared a great deal with Gaizauskas's models of why beings are given rights.

For Prescott, the self has many parts: physical, social, temporal, conceptual and private. These parts of the self develop in different ways at different times and at different rates. At Sheffield Robotics, Prescott leads a project that is attempting to replicate these developments in an iCub robot, by giving the robot a physical self (an awareness of its own body and the space around it) and a temporal self, not just a sense of its own past history, but an ability to imagine itself in the future. The question as to where and/or when we might say that artificial intelligence is (self-)aware is not the only, or even the most important, criteria upon which we can judge whether artificial intelligence can be deemed to be a 'person' as, for Prescott (as for Gaizauskas) there are many other criteria that would qualify robots and artificial intelligence as worth for ethical consideration, or of rights and obligations.

Following the presentations, there followed a very lively discussion, involving the presenters and the audience. It was noted that the papers represented five very different approaches, and the contributions from the audience added even more diverse voices to the conversation, which showed a terrific breadth of views and demonstrated that such a plurality of voices and contributions will be required to make real progress on the question of ethics and effective governance with robots and artificial intelligence. There was a great deal of discussion on the questions of personhood, or 'e-personhood', and whether robots and artificial intelligence qualify for ethical consideration, and on the expanded notions of selfhood and the basic qualities necessary for rights and obligations offered by Ponce Del Castillo, Gaizauskas and Prescott. A good part of the discussion also focussed on the public understandings, or misunderstandings, of robots and artificial intelligence — unsurprising as the Science in Public conference has a keen interest in science communication — and, following on from my talk and Penn's, the role the popular media and historical contextualisation can play in improving our public conceptualisations and policy on future technologies. From these discussions, we have begun to make some hopefully enduring collaborations that will seek to unite these different approaches in analyses of future legislation, ethics and popular representations of robots and artificial intelligence.

Author

Michael Szollosy is a Research Fellow in Sheffield Robotics at the University of Sheffield, looking at the cultural influences and social impacts of robots, AI, VR and other emerging technologies. He is Sheffield Robotics' lead for public engagement and responsible research. E-mail: m.szollosy@sheffield.ac.uk.

How to cite

Szollosy, M. (2017). 'Robots, AI, and the question of 'e-persons' — a panel at the 2017 Science in Public conference, 10–12 July 2017'. *JCOM* 16 (04), C05.



This article is licensed under the terms of the Creative Commons Attribution - NonCommercial - NoDerivativeWorks 4.0 License. ISSN 1824-2049. Published by SISSA Medialab. jcom.sissa.it