eLaw
Working Paper Series

No 2019/002 - ELAW– 24 April 2019

Responsibility in Robot and AI Environments

Eduard Fosch Villaronga

Universiteit Leiden
eLaw

Discover the world at Leiden University

Responsibility in robot and AI environments*

E. Fosch-Villaronga**

"When we want to do something while unconsciously certain to fail,
we seek advice so we can blame someone else for the failure"
Nassim Nicholas Taleb.

Abstract: Unintended harm can occur in the course of operation of a robot: a robot grasper can hit a person, a user can fall when wearing a robotic exoskeleton, and the malfunctioning of the robot operating system can erroneously hit a glass wall thinking there was no barrier. Moreover, harm can also appear after a while, after having used the robot continuously for some time. Since robots are systems constructed with different components, including hardware, software, and cloud services that may be provided by different companies, and they perform tasks with a degree of autonomy, it might be uncertain who is responsible if something goes wrong. In this article, I address responsibility issues arising from the use and development of robot and AI technologies. The first two sections explain the complex ecosystem of robot and AI technologies, and map different sources of liability in such environments. The third and fourth sections describe what the 'responsibility gap' is and compile different theories extending the responsibility to non-human agents. In the end, since multiple parties may exercise various types of control over the creation and development of a robot, I suggest that technological complexity should be no reason to remove human responsibility in the context of robots and AI.

Keywords: responsibility, accountability, liability, robot, artificial intelligence, AI, safety, agenthood.

1. Introduction

Unintended harm can occur in the course of operation of a robot: a robot grasper can hit a person, a user can fall when wearing a robotic exoskeleton, and the malfunctioning of the robot operating system can erroneously hit a glass wall thinking there was no barrier. Moreover, harm can also appear after a while, after having used the robot continuously for some time. In the case of lower-limb exoskeletons, for instance, it could well be that the users' muscles are activated, but user's do not detect whether it is part of the normal robot usage, or not. Some of the users of this technology might lack the capacity to feel the legs, or may merely not know how their muscles were activated before they had the injury. In a recent study, rehabilitation patients could not provide reliable feedback to physicians or therapists because they lacked the means on how to do so. [1]

* This article was funded by the LEaDing Fellows Marie Curie COFUND fellowship, a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 707404.
** Dr. Eduard Fosch-Villaronga is a Marie Skłodowska-Curie Postdoctoral Researcher at the eLaw Center for Law and Digital Technologies at Leiden University in the Netherlands.
[1] E. Datteri, 'Predicting the long-term effects of human-robot interaction: A reflection on responsibility in medical robots', *Science and engineering ethics* (19) 2013, issue 1, p. 139-160 (hereinafter: Datteri 2013).

Retrospective liability should apply if there is a causal link between the robotic device and the future harm, Datteri argues.[2]

Responsibility, liability, and accountability are different concepts. In the majority of the Latin languages, for instance, these three concepts are expressed through the same word, for instance, *responsabilitat* (Catalan) *responsabilité* (French), *responsabilità* (Italian), *responsabilidad* (Spanish), *responsabilidade* (Portuguese), *responsabilitate* (Romanian). It is true that in Catalan, Spanish and French the word 'accountability' in some cases could be translated slightly different (*rendició de comptes*, *rendición de cuentas*, *rendition des comptes*), as well as in Romanian the word liability can be translated into *răspundere*; but, generally speaking, they are used interchangeably.

Black's Law Dictionary defines *responsibility* as the 'obligation to answer for an act done, and to repair any injury it may have caused.'[3] This definition refers to the state of having a duty to deal with something or someone, like a mother for her child. Dworkin argues that it is to act with due diligence and that non-compliance could lead society to claim for any consequences flowing from the non-compliance. When this responsibility is a legal requirement, that is to say when someone is bound or obliged by the law (in other words, legal responsibility), this is called liability. Black's Law Dictionary defines this as 'the state of being bound or obliged in law or justice to do, pay, or make good something; legal responsibility.'[4] Accountability is required or expected to justify actions or decisions, and normally refers to the party that must report activities and take responsibility for them.[5] As explained in Black's Law Dictionary, this is to keep the party honest and responsible. Kool straightforwardly summarizes this threesome: 'accountability follows responsibility, whereas accountability aims at establishing liability.'[6] As she points out, however, liability and accountability are not synonymous with responsibility.

However, what do all these concepts mean for robots and artificial intelligence? Robots are systems constructed with different components, including hardware, software, and cloud services that may be provided by different companies, and they perform tasks with a degree of autonomy; it might be uncertain who is responsible if something goes wrong.[7] In this article, I shed the light on this, and investigate responsibility issues arising from the use and development of robot and AI technologies.

In the first two sections, I explain the complex ecosystem of robot and AI technologies, and map different sources of liability in such environments. In the third and fourth sections, I describe what the 'responsibility gap' is and compile different theories extending the responsibility to non-human agents. In the end, since multiple parties may exercise various types of control over the creation and development of a robot, I suggest that technological complexity should be no reason to remove human responsibility in the context of robots and AI.

---

[2] Datteri 2013.

[3] 'What is RESPONSIBILITY?', *Black's Law Dictionary* s.d., thelawdictionary.org/responsibility.

[4] 'What is LIABILITY?', *Black's Law Dictionary* s.d., thelawdictionary.org/liability.

[5] 'What is ACCOUNTABILITY?', *Black's Law Dictionary* s.d., thelawdictionary.org/accountability.

[6] R.S.B. Kool, '(Crime) Victims' Compensation: The Emergence of Convergence', *Utrecht L. Rev.* (10) 2014, issue 3, p. 14-26.

[7] E. Fosch-Villaronga & C. Millard, 'Cloud Robotics Law and Regulation', *Queen Mary School of Law Legal Studies Research Paper* No. 295/2018, available at *SSRN*, ssrn.com/abstract=3305353 (hereinafter: Fosch-Villaronga & Millard 2018).

2. Robots and AI usually involve complex ecosystems

2.1 In theory

The primary responsibility for product safety falls on producers, which may include 1) manufacturers that supply products on the market, 2) manufacturers that provide products for own use, 3) importers (in some instances), and 4) businesses/users that change features of the product in order to customize it.[8]

For the Robolaw[9] project, the introduction of robots in society raises questions concerning the potential damage to a primary or secondary user, or to the equipment. They argue that the development risk defense for product liability should be considered under Article 7(e) of the European Union Council Directive 85/374/EEC on the approximation of the laws,[10] regulations and administrative provisions of the Member States concerning liability for defective products. According to this article: 'The producer shall not be liable as a result of this Directive if he proves … that the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered.'

If a robot is designed to perform a task and runs proprietary software, then liability may be quite straightforward. As Calo explains,[11] if a Roomba harms someone when vacuuming the floor, then iRobot will probably be held liable because it has built the hardware and written the software.[12] If, however, a Roomba is modified or is used for other purposes, then iRobot might seek to avoid liability.

2.2 In practice

However, robots using cloud services might be considered a hybrid category between product and (interconnected) service(s), bringing about legal uncertainty with regards to the application of such product liability framework.[13] Indeed, robots are systems constructed with different components, including hardware, software, and cloud services that may be provided by different companies,[14] and it might not be straightforward to determine who is in control at any time in the course of operation of a robot. Buddy, for instance, is an open source and 'open-hardware-botic' robot – meaning anyone can build accessories for the physical robot embodiment.[15]

---

[8] F. Stoica, 'EU product safety framework for advanced robots & autonomous systems', European Commission s.d., http://ec.europa.eu/information_society/newsroom/image/document/2017-30/felicia_stoica_-_the_existing_eu_safety_framework_with_regard_to_autonomous_systems_and_advanced_robots__iot-systems_6210B836-9707-D592-D33613EE1C6F086A_46145.pdf.

[9] E. Palmerini and others, 'RoboLaw. Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics', *RoboLaw* 22 September 2014, robolaw.eu/RoboLaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf.

[10] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products (*OJ* 1985, L 210/29).

[11] R. Calo, 'The Need to Be Open: U.S. Laws Are Killing the Future of Robotics', *Mashable* 1 January 2014, mashable.com/2014/01/01/us-law-robotics-future.

[12] J. McCurry, 'South Korean woman's hair 'eaten' by robot vacuum cleaner as she slept', *The Guardian* 9 February 2015.

[13] Fosch-Villaronga & Millard 2018.

[14] G.N. La Diega & I. Walden, 'Contracting for the 'Internet of Things': Looking into the Nest', *EJLT* (7) 2016, issue 2.

[15] 'Buddy The Emotional Robot', *Bluefrog Robotics* 2018, buddytherobot.com/en/buddy-the-emotional-robot.

Determining which party should be held accountable for particular harm is particularly challenging for issues that arise from the composition of, and interactions between, components managed by different entities, rather than from a single entity failing to act appropriately.[16] In this type of ecosystems, chains of responsibility may be extensive, complex and often opaque, even to the contracting parties.[17] Indeed, some users of those services report that, on many occasions, 'you only know what your provider tells you.'[18] Robot providers may be dependent on the different cloud service provider(s) and sub-providers with non-negotiable terms and conditions, and it may be challenging to understand faults or errors, determine causality, and attribute responsibility in such complex ecosystems.

In the case of damage resulting from a technology failure in a cloud robotics ecosystem, an aggrieved party may sue a robot provider for compensation for a failure to avoid and manage foreseeable risks.[19] However, to establish an intelligible causation, a claimant may need to understand how the cloud robotics ecosystem works in general and access to the cloud provider's information regarding the specific incident. In such cases, it may be difficult to have a legal basis to compel the cloud service providers to explain how their system works without compromising trade secret information.[20] This complex, multi-party, ecosystem may complicate the process of attributing legal responsibility in robotics and AI environments.

The parties may seek to establish their responsibilities through contracts beforehand to prevent them from having to deal with such issues after harm occurs. However, most terms and conditions are standard, non-negotiable, and this might affect how the liability.[21] Robot providers should consider whether these standard terms of service are appropriate, given the particular risks of real-world, physical effects. Alternatively, they could always try to negotiate specific terms, although it is unlikely.

The European Commission (the Commission) warns that more legal certainty is needed for the liability for service failures, user rights protection in unilaterally and automatically updated systems, the ownership of data created through these applications, and how disputes will be resolved.[22] Although the Commission highlighted these issues for the Internet of Things, robots are a *thing* that may well benefit from the use of the infrastructure of Internet of Things.

3. Sources of liability

A robot may behave strangely because the designer has failed to take into account certain environmental variables. It could be that the particular evaluation method used was not correct, or because the roboticists did not want to be more cost efficient when training the robot. It could also

---

[16] J. Singh and others, 'Accountability in the IoT: Systems, Law and Ways Forward', *Computer* (51) 2018, issue 7, p. 54-65 (hereinafter: Singh and others 2018).

[17] C. Millard, W.K. Hon & J. Singh, 'Internet of Things Ecosystems: Unpacking Legal Relationships and Liabilities', in: *2017 IEEE International Conference on Cloud Engineering (IC2E)*, Piscataway, New Jersey: IEEE 2017, p. 286-291; Fosch-Villaronga & Milllard 2018.

[18] B. Khoe, 'Serverless IoT at iRobot', *InfoQ* 30 March 2018, infoq.com/presentations/serverless-iot-irobot.

[19] Singh and others 2018.

[20] Singh and others 2018.

[21] C. Millard, 'Cloud Computing Transactions', in: C. Millard (ed.), *Cloud Computing Law*, Oxford: Oxford University Press 2013, p. 37.

[22] Unleashing the Potential of Cloud Computing in Europe, COM (2012) 529 final.

be because there have been wrong extrapolations from limited samples, or because insufficient training data have been used, all of which affected the learning model of the robot negatively.[23]

These humanly taken decisions over data analysis for robot task learning and performance may compromise the correct functioning of a robot.[24] Apart from discriminatory scenarios, such errors in data analysis in the context of physical robots designed to work autonomously can have direct, and in some cases severe, physical consequences. Indeed, 'systems that simply output a recommendation to human users, such as speech systems, typically have relatively limited potential to cause harm. By contrast, systems that exert direct control over the world, such as machines controlling industrial processes, can cause harms in a way that humans cannot necessarily correct or oversee.'[25]

Robots may also lose functionality because of reduced connectivity, challenging the safety and the security of the operation. In such cases, it will be challenging to identify the origin of a problem, the calculation of its consequences and its subsequent future impacts.[26] If a robot causes unintended harmful behavior,[27] it may be necessary to examine the terms of the relevant contracts between the different parties to determine liability.

A robot impact assessment may indicate that specific critical functions must operate onboard the robot and independent of any connectivity, or even lead to the conclusion that robots cannot perform specific critical tasks due to risks of inappropriate function or inoperability that might result from connection loss.[28] This assessment could be informative to the relevant authorities (for law enforcement) but also to users that want to know how to use a robot safely.

4. The responsibility gap

An unintended harmful behavior may occur when, despite having a well-defined function, a robot behaves in a way that differs from the designer's intent.[29] Many researchers support the idea that, if a robot learns as it operates, and the robot itself can, in the course of its operation, change the rules by which it acts, then there is no reason why humans should be held responsible for the autonomous behaviors of such a robot.[30]

---

[23] R. Richardson, J. Schultz & K. Crawford, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice', *NYU L. Rev. Online* (forthcoming), available at *SSRN* 13 February 2019, papers.ssrn.com/abstract =3333423 (hereinafter: Richardson, Schultz and Crawford 2019).

[24] Richardson, Schultz and Crawford 2019.

[25] D. Amodei and others, 'Concrete problems in AI safety', *arXiv* 21 June 2016, *https://arxiv.org/pdf/*1606.06565.

[26] Y.H. Weng & S.T.H. Zhao, 'The legal challenges of networked robotics: From the safety intelligence perspective', in: *International Workshop on AI Approaches to the Complexity of Legal System*, Berlin/Heidelberg: Springer 2011, p. 61-72.

[27] F. Pistono & R.V. Yampolskiy, 'Unethical research: How to create a malevolent artificial intelligence', *arXiv* 1 September 2016, arxiv.org/abs/1605.02817 (hereinafter: Pistono & Yampolskiy 2016).

[28] E. Fosch-Villaronga, 'Creation of a Care Robot Impact Assessment', *WASET. International Science Journal of Social, Behavioral, Educational, Economic and Management Engineering* (9) 2015, issue 6, p. 1817-1821.

[29] Pistono & Yampolskiy 2016.

[30] A. Matthias, 'The responsibility gap: Ascribing responsibility for the actions of learning automata', *Ethics and information technology* (6) 2004, issue 3, p. 175-183; T. Hellström, 'On the moral responsibility of military robots', *Ethics and information technology* (15) 2013, issue 2, p. 99-107.

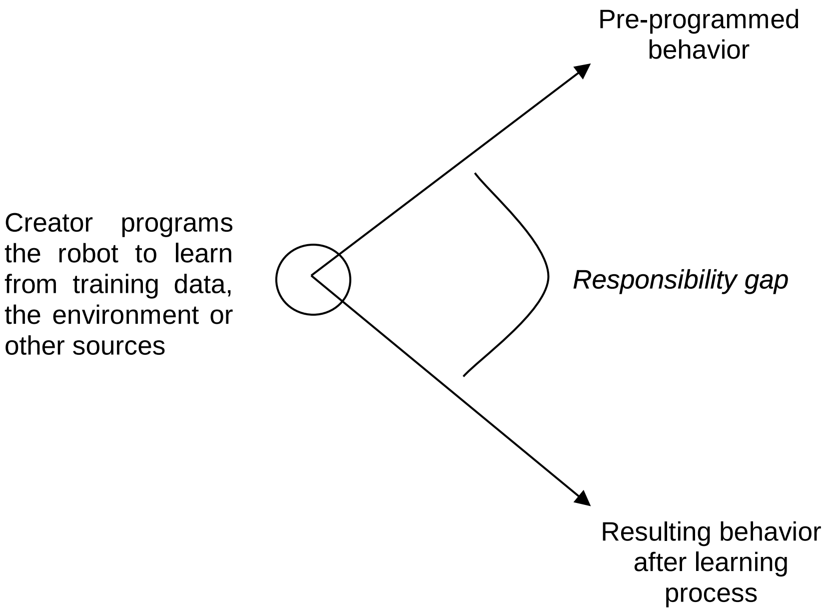

Fig. 1 Graphic explanation of the responsibility gap

Johnson argues, however, that the concept of responsibility within the robotic domain was still underdeveloped and that these dominant positions on this domain do generally not analyze deeply the real meaning of responsibility of an artificial agent. She believes that in reality, the responsibility gap challenges neither responsibility nor liability concepts, but rather accountability.[31] With the growing use of impact assessments in the European Union, it is evident that the creation of autonomous robotic technology has considerably challenged responsibility.

This analysis persuaded the European Parliament, who called on the Commission to create a specific legal status for robots – in the form of electronic persons - and to hold them responsible when causing damage.[32] The expert opinion differs between those who support the idea of robot personality,[33] and many others who are very concerned about it.[34]

The fact that the robot behaves in a different way to the designer's intention should not necessarily exempt the designer from being responsible. Robots are complex systems comprising various components that may entail multiple processes over which different persons, natural or legal,

[31] D.G. Johnson, 'Technology with no human responsibility?', *Journal of Business Ethics* (127) 2015, issue 4, p. 707-715 (hereinafter: Johnson 2015).

[32] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), A8-0005/2017.

[33] M. Spiliopoulou-Kaparia, 'The evaluation of Directive 85/374/EEC on liability for defective products and Directive 2006/42/EC on machinery. European Stakeholder Forum – Workshop on Regulatory challenges for a digitizing industry Essen', *European Commission* 1 February 2017, ec.europa.eu/futurium/en/system/files/ged/b3-spiliopoulou-liability.pdf.

[34] Opinion of the European Economic and Social Committee on 'Artificial intelligence — The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society' (*OJ* 2017, C 288/1); 'Report of COMEST on robotics ethics', SHS/YES/COMEST-10/17/2 Rev (14 September 2017), unesdoc.unesco.org/ark:/48223/pf0000253952; 'Open letter to the European Commission: Artificial Intelligence and robots', *Robotics Openletter* s.d., robotics-openletter.eu.

exercise control and, therefore, might be held responsible.[35] Technological complexity should not, in itself and therefore, be a reason for removing liability which might otherwise arise.[36]

5. Who is responsible, then?

There are authors that argued that the more autonomous electronic agents become, the more lawyers should accept that minor changes to the law cannot be the solution.[37] Part of the legal community believes that if agents (robots in this case) are capable of deciding on their own, they should deserve some legal status, i.e., a status recognized in the law for non-humans.

This has led a lot of legal scholars to reflect on which metaphor or theory best explains the extension of responsibility to non-human agents:

- Chopra and White reflected on the possibility to use the electronic agent metaphor and stated that the U.S. Uniform Electronic Transactions Act (1999) already contemplated the possibility of an automated transaction formed by the interaction of electronic agents of the parties.[38]
- Coeckelbergh argued that maybe the animal metaphor could be used, as animals were considered until very recently things - now considered *sentient beings*.[39]
- Mady Delvaux resonated that robots are not humans, and will never be, but that this is similar to corporations, which are legal persons and enjoy some of the rights of natural persons without being human.[40]
- Pagallo traveled back in time and explained in his book that, in Roman law, there was this institution called *peculium* for those transactions made by the slave, at that moment considered a thing, in the name of the master. He argues that the robot and the responsibility could be inspired by this institution and be called *digital peculium*.[41]

In this respect, the European Parliament does believe in the existence of a responsibility gap. It clearly states 'the ability to learn from experience and take independent decisions – has made them [robots] more and more similar to agents that interact with their environment and can alter it significantly; whereas, in such a context, the legal responsibility arising from a robot's harmful action becomes a crucial issue.'[42] The Parliament also admits, however, that without any further change, in the current legislative framework, it would be impossible to hold robots responsible.

---

[35] J. Singh and others, 'Responsibility & Machine Learning: Part of a Process', *SSRN* 28 October 2016, papers.ssrn.com/ abstract=2860048.

[36] Johnson 2015; Fosch-Villaronga & Millard 2018.

[37] S. Chopra & L. White, 'Artificial agents-personhood in law and philosophy', in: *Proceedings of the 16th European Conference on Artificial Intelligence*, Amsterdam: IOS Press 2004, p. 635-639 (hereinafter: Chopra & White 2004).

[38] Chopra & White 2004. See https://rightsignature.com/legality/ueta-act.

[39] M. Coeckelbergh, 'Humans, animals and robots. A phenomenological approach to human-robot relations', *International Journal of Social Robotics* (3) 2011, issue 2, p. 197-204; 'Animal welfare', *European Commission* s.d., ec.europa.eu/food/animals/welfare_en.

[40] J. Vincent, 'Giving robots 'personhood' is actually about making corporations accountable', *The Verge* 19 January 2017, theverge.com/2017/1/19/14322334/robot-electronic-persons-eu-report-liability-civil-suits.

[41] U. Pagallo, *The laws of robots: crimes, contracts and torts* (vol. 10), Berlin: Springer Science & Business Media.

[42] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), A8-0005/2017.

Other authors argue that responsibility strictly relates to self-consciousness and that only if future robotic technologies could develop this characteristic might they be granted personhood.[43] Still, legal doctrine is not clear on how to deal with the new actor status. Teubner believes that Law has already welcomed new agents in the legal system in the past, e.g., electronic agents and animals and that depending on the chosen doctrinal construction, the legal consequences (protection of the agent, responsibility for the behavior of the agent) would be different.[44]

Similarly, Laukyte[45] argues that we have already witnessed the extension of rights to other entities such as animals (through the theory of self-consciousness or as Darling explained through the theory of anthropomorphism[46]), and corporations (like the free speech of corporations). Laukyte argues that there is no reason why in the future we could not extend rights to robots, and so does Darling. Following the theory of self-consciousness, Laukyte reports an example of the work of Gold and Scassellati.[47] These engineers created a humanoid robot with self-awareness capabilities similar to humans, in fact – already in 2007 – this robot was able to recognize itself in a mirror. Laukyte also mentions the theory of pain, that is, we protect animals because they feel pain.[48] Although seemingly impossible, there are already researchers trying to teach robots how to handle and recognize pain. Could we think that in the future robots that can feel pain, also have a recognized agenthood?

Although Laukyte and Darling seem to be in favor of the extension of legal rights to robots (although from different theories), this appears to be from the defensive institution's point of view: robots need to be protected from abuse. Indeed, Teubner argues that when society has identified animals as new agents, it has done that to protect them against injuries. Teubner believes, on the contrary, that electronic agents (and by extension robots) create an aggressive, productive institution from which society needs to be protected. In other words, Teubner believes that society needs to be protected from these new agents instead of protecting them.[49]

Johnson uses a very futuristic scenario to talk about whether the society will accept technology with no human responsibility or not. She argues that maybe in the future machines will have to explain what they did wrong and why they did so. This scenario might not be that far away from reality: one of the most significant problems that current roboticists have is how to trace back the behavior of a robot. In other words, sometimes the robot behaves, and roboticists do not know precisely why. In the future, not only the use of black boxes but also the use of explanatory arguments from the machine could be beneficial for users and judges to know what went wrong.

---

[43] E.J. Koops, M. Hildebrandt & D.O. Jaquet-Chiffele, 'Bridging the accountability gap: Rights for new entities in the information society', *Minn. J. L. Sci. & Tech* (11) 2010, issue 2, p. 497-561.

[44] G. Teubner, 'Rights of non-humans? Electronic agents and animals as new actors in politics and law', *Journal of Law and Society* (33) 2006, issue 4, p. 497-521 (hereinafter: Teubner 2006).

[45] M. Laukyte, 'The Capabilities Approach as a Bridge Between Animals and Robots', EUI MWP 2013/05, available at http://cadmus.eui.eu/handle/1814/27058 (hereinafter: Laukyte 2013).

[46] K. Darling, 'Extending legal protection to social robots: The effects of anthropomorphism, empathy and violent behavior towards robotic objects', in: R. Calo, A.M. Froomkin & I. Kerr (eds.), *Robot law*, Cheltenham: Edward Elgar Publishing 2016, p. 213-231 (hereinafter: Darling 2016).

[47] K. Gold & B. Scassellati, 'A Bayesian Robot That Distinguishes "Self" from "Other"', in: *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2007, p. 29.

[48] Laukyte 2013.

[49] Teubner 2006.

6. Human responsibility

Other authors' viewpoints clash with the responsibility gap.[50] These opinions usually are in line with the strict liability regime, that is, when a person/entity is liable regardless of the culpability element; or with the existence of professional codes of conduct, values and ethical standards that establish a "professional responsibility" that would prevent roboticists from creating uncontrollable robots.[51]

The Engineering and Physical Sciences Research Council (EPSRC) already said plainly that the responsible agents are the humans and not the robots.[52] Indeed, the creation of new technologies involves 1) persons that decide on the nature of these technologies and 2) people that accept these decisions.[53]

According to Johnson's assumption, the responsibility gap refers to the scenario where the persons involved in the creation of a robot have decided to create a technology that is uncontrollable, and that the society (consumers especially) has accepted the fact that there is no human responsibility for this uncontrollable robot.[54] However, people would not accept machine liability in the case of uncontrollable machines.[55] When technology acts in an unexpected way, e.g. when the Google photos' algorithm label black people with gorillas or when YouTube Kids contains videos with inappropriate content for young children, the first thing to wonder should be who is to blame.[56] As Johnson explains, people not only expect that certain behaviors will not happen again but certainly ask for something to be done to ensure it. In other words, people want to know who is accountable. [57]

Johnson admits that, because robots are typically built and used by many people it might be challenging to know in the future who should be accountable for them. In fact, and taking into account the second scenario above-mentioned (i.e., people accept these decisions), she believes that future generations will have two choices: either accept technologies with no human responsibility or attribute accountability to each of the contributors of the robot chain. In other words, 'the responsibility gap depends on human choices, not on the complexity of the technology.'[58]

7. Conclusion

As the same as what happens with the use of robots in highly unstructured environments and diverse scenarios, 'only the diffusion and real use of the device – and subsequent accidents

[50] M. Santoro, D. Marino & G. Tamburrini, 'Learning robots interacting with humans: from epistemic risk to responsibility', *AI & Society* (22) 2008, issue 3, p. 301-314 (hereinafter: Santoro, Marino & Tamburrini 2008); M. Nagenborg and others, 'Ethical regulations on robotics in Europe', *AI & Society* (22) 2008, issue 3, p. 349-366 (hereinafter: Nagenborg and others 2008); J.J. Bryson, M.E. Diamantis & T.D. Grant, 'Of, for, and by the people: the legal lacuna of synthetic persons', *Artificial Intelligence and Law* (25) 2017, issue 3, p. 273-291.
[51] Nagenborg and others 2008.
[52] 'Principles of robotics', *EPSRC* September 2010, epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics.
[53] Johnson 2015.
[54] Johnson 2015.
[55] Santoro, Marino and Tamburrini 2008.
[56] A. Chander, 'The racist algorithm?', *Mich. L. Rev.* (115) 2017, issue 6, p. 1023-1045.
[57] Johnson 2015.
[58] Johnson 2015.

caused – will provide more reliable data.'[59] However, should society allow the occurrence of these accidents to have the actual data? It does not seem to work the same way with other technologies. Airplanes have a clear regulation on simulator hours for pilot training and a protocol before take-off for security purposes. At the moment, however, robots may apply forces that could be destructive and it is not clear what protocol should be applied to them.

More qualitative and quantitative data is needed to understand the likelihood of occurrence (and the extension of the damage) of harms after robot usage, and whether some extra safeguards should be implemented in this respect. There are authors that support the idea of extending legal personality to robots, mainly to provide a mechanism for applying directly to robots various obligations that currently apply only to individuals and legal persons such as companies. However, I am not sure 'whether it is ethically acceptable to deploy some robotic system or technology for tasks that involve (potentially harmful) human-robot interactions'[60] nor I am convinced that holding the robots responsible will solve many of the arisen responsibility issues.

[59] A. Bertolini and others, 'On robots and insurance', *International Journal of Social Robotics* (8) 2016, issue 3, p. 381-391.
[60] Datteri 2013.