THE COMPLEXITY OF CRIMINAL LIABILITY OF AI SYSTEMS[*]

by

NORA OSMANI[**]

Technology is advancing at a rapid pace. As we anticipate a rapid increase in artificial intelligence (AI), we may soon find ourselves dealing with fully autonomous technology with the capacity to cause harm and injuries. What then? Who is going to be held accountable if AI systems harm us?

Currently there is no answer to this question and the existing regulatory framework falls short in addressing the accountability regime of autonomous systems. This paper analyses criminal liability of AI systems, evaluated under the existing rules of criminal law. It highlights the social and legal implications of the current criminal liability regime as it is applied to the complex nature of industrial robots. Finally, the paper explores whether corporate liability is a viable option and what legal standards are possible for imposing criminal liability on the companies who deploy AI systems.

The paper reveals that traditional criminal law and legal theory are not well positioned to answer the questions at hand, as there are many practical problems that require further evaluation. I have demonstrated that with the development of AI, more questions will surface and legal frameworks will inevitably need to adapt. The conclusions of this paper could be the basis for further research.

KEY WORDS

Artificial Intelligence, Autonomous Systems, Criminal Law, Criminal Liability

[**] osmani.nora@hotmail.com, PhD Candidate, Faculty of Law, Ss. Cyril and Methodius University, Skopje, Republic of North Macedonia; visiting researcher at Masaryk University, Brno, The Czech Republic.

1. INTRODUCTION

Robotics and AI are moving towards a Cambrian explosion,[1] given the explosive and rapid developments in the field. A lot of countries have recognized the enormous importance of the AI industry[2] and are now actively competing for the global AI market.[3] In terms of revolution, machine learning created the ability for AI entities to accumulate experience and learn from it.[4] During an experiment, researchers from the *Georgia Tech Institute* fielded two autonomous aircraft (auto-piloted airplanes), which shared and communicated information to each other using data communication software and Wi-Fi systems; thus no human control was involved whatsoever.[5]

The self-development ability of AI systems presents society with machines that have their own needs and goals, referred to in the literature as *Artificial Super Intelligence (ASI)*.[6] These thinking machines[7] can act directly upon their environment,[8] thus raising questions about liability for

---

[1] Pratt, GA. (2015) Is a Cambrian Explosion Coming for Robotics?. *Journal of Economic Perspectives*, 29 (3), p. 55.

[2] Countries such as: Germany, Switzerland, Canada, Hong Kong, Singapore, South Korea, France, United Arab Emirates, Japan, Russia, and Israel aspire to participate in the global AI market and are actively supporting the development of the AI industry through their tech--hubs. For more information, see: Deep Knowledge Analytics. (2018) *Artificial Intelligence Industry in the UK*. [online] London: DKA. Available from: https://www.dka.global/ai-in-uk-report [Accessed 12 May 2019].

[3] China and the US remain the leaders of the AI race, but they are actively competing to outperform each other and to become leaders in the global IT market. Chinese President *Xi Jinping* has committed $150 billion in government funding in order to make China the undisputed global leader in the AI race by 2030. Ibid.

[4] Machine learning is described as an inductive method of learning in which the AI system analyses various given data and identifies patterns for future use without being explicitly programmed for that purpose. In this sense, the computer's behaviour is not predictable by either the operator–owner or the original programmers. See, e.g., Vojislav, K. (2001) *Learning and Soft Computing: Support Vector Machines, Neutral Networks and Fuzzy Logic Models*. Cambridge: The MIT Press; Karnow, C.E.A. (2013) The Application of Traditional Tort Theory to Embodied Machine Intelligence. In: Ryan Calo, Michael Froomkin, Ian Kerr (eds.). *Robot Law*. Cheltenham: Edward Elgar Publishing Ltd, pp. 51–77.

[5] Toon, J. (2017) *Swarms of Autonomous Aerial Vehicles Test New Dogfighting Skills*. [online] Atlanta: GeorgiaTech. Available from: http://www.rh.gatech.edu/news/590743/swarms-autonomous-aerial-vehicles-test-new-dogfighting-skills [Accessed 15 May 2019].

[6] Radutniy, O.E. (2017) Criminal Liability of the Artificial Intelligence. *Problems of Legality*, (138), p. 136.

[7] Hallevy, G. (2010) 'I, Robot – I, Criminal' – When Science Fiction Becomes Reality: Legal Liability of AI Robots Committing Criminal Offenses. *Syracuse Science and Technology Law Reporter*, 22 (Spring), p. 1.

[8] In his research, *Calo* defines robots as machines with three qualities: *"(1) a robot can sense its environment, (2) a robot has the capacity to process the information it senses, and (3) a robot is organized to act directly upon its environment"*. See: Calo, R. (2016) *Robots in American Law. University of Washington School of Law Research Paper No. 2016-042*. Available from: https://ssrn.com/abstract=2737598 [Accessed 22 June 2019].

their actions. While they deliver great social benefits, can society's interaction with robots also cause harm?

In an increasingly automated world, these threats are becoming real and many people have spoken out about their fears. At a tech conference that took place on 6 November, 2017 in Lisbon, Portugal, the physicist *Dr. Stephen Hawking* warned us about the dangers of AI, stating that AI can be the worst case of human intelligence.[9] The entrepreneur *Elon Musk* has also expressed his concerns. He calls for establishment of a regulatory authority that would monitor development of AI, citing it as the most likely cause of World War III.[10]

The fact that the *US Air Force (USAF)* already uses some semi- and fully autonomous technology and expects to increase these systems' autonomy in the future,[11] gives credence to these fears. According to a study conducted by researchers from *Oxford* and *Yale Universities*, AI systems will outperform humans in many activities in the next ten years.[12]

In our coexistence with intelligent agents, one can naturally ask: *Who should be sued in court if a robot (that is, only fully autonomous robots, that operate without human control) makes a mistake, thereby killing, or causing serious injuries?* These questions have been puzzling academics and researchers. Even though this problem has been addressed before,[13] only recently it has attracted enough attention for a prompt examination. Due to the absence of direct legal regulation of AI, the questions at hand still remain unclear. The uncertainty largely stems from the complex nature

---

[9]   See, for example: Murphy, M. (2017) *Stephen Hawking: AI Could Be Best – or Worst – Thing in Human History*. [online] New York City: MarketWatch. Available from: https://www.marketwatch.com/story/stephen-hawking-ai-could-be-best-or-worst-thing-in-human-history-2017-11-06 [Accessed 29 October 2018].

[10] See, for example: Osborne, S. (2017) Elon Musk Calls for Urgent Laws on Robot as They Will Soon Be Risk to Public. *Express*, 28 November. [online] Available from: https://www.express.co.uk/news/science/885344/elon-musk-artificial-intelligence-robotics-regulation [Accessed 29 October 2018].

[11] Palmer, A. (2010) Autonomous UAS: a partial solution to America's future airpower needs. [pre-print] Submitted to: *Air University in partial fulfillment of the graduation requirements*. Available from: https://apps.dtic.mil/dtic/tr/fulltext/u2/1018416.pdf [Accessed 29 November 2018].

[12] Grace, K. et al. (2018) Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts. *Journal of Artificial Intelligence Research*, 62, pp. 729–754.

[13] See, for example: Willick, M. (1983) Artificial Intelligence: Some Legal Approaches and Implications. *AI Magazine*, 4 (2), p. 5. Available from: https://aaai.org/ojs/index.php/aimagazine/article/view/392 [Accessed 29 November 2018]; Karnow, C.E.A. (1996) Liability for Distributed Artificial Intelligences. *Berkeley Technology Law Journal*, 11 (1), pp. 147–204. Available from: https://lawcat.berkeley.edu/record/1115611?ln=en [Accessed 29 November 2018].

of industrial robots and the multiple parties involved in the process of their manufacture.

The aim of this research is to outline a comprehensive analysis of significant liability models that address the responsibility for hazards caused by autonomous systems. This paper does not intend to offer a detailed answer for the models presented, considering that such an analysis goes beyond a research paper. Therefore, theories and explanations intend to serve the reader as guideposts to further exploration.

To limit the scope of the study, the paper primarily reflects a criminal approach to liability, and the focus is mostly in the assessment of fault rather than in the allocation of cost. Thus, a tort law approach to providing compensation for damage or physical injuries caused by AI systems is not an object of analysis in this article.

The paper consists of three main parts. The first part examines direct liability of AI systems. The matter is approached by considering that, in the near future, robots might become sophisticated machines having human-like abilities. The paper examines whether AI agents could be granted legal personhood and whether they can be held criminally liable for their actions. The second part examines whether responsibility can be allocated to different agents in the production chain, i.e. imposing strict liability on operators of the AI agents, such as programmers, designers, owners, and other parties involved in the process of manufacturing. Finally, corporate criminal liability is considered as a plausible method to address the wrongdoings of autonomous systems.

The article argues that current criminal law and the current legal system are not well positioned to balance a concern for physical safety with incentivizing innovation of AI and industrial robots. The doctrine of direct liability of AI systems is problematic from a normative perspective. In the absence of a regulatory framework governing the legal status of AI systems, the question of whether AI systems could become legal persons remains only theoretical. The concept of AI as a tool and strict liability of manufacturers is also problematic, as the plaintiff may find it very difficult to prove that the AI system was defective in the moment it left the manufacturer or its developer. On the other hand, there are no clear grounds on which to allocate liability of producers based on negligence. With regard to AI systems as self-learning entities, it is hard to predict their behaviours and define the cause-effect pattern; therefore, the operators

of AI systems may not be in a favourable position to foresee risks and provide due diligence. Given the fact that AI represents a great power in the hands of giant tech companies, it may fairly be said that corporate liability should gain more attention from researchers, academics, and policymakers. However, there is a need for an international regulatory response that would set down basic rules that respond to the factual problem.

2. DIRECT CRIMINAL LIABILITY OF AI SYSTEMS

Are machines legal authors of their actions? The first question is whether or not we can recognize AI agents as legally responsible for their behaviours in terms of traditional criminal law theories.

In criminal law, to impose criminal liability for intentional offences, two main criteria must be met.[14] One is the factual element (*actus reus*), which contains the criminal conduct, and the other one is the mental element (*mens rea*), which consists of the general intent of the offender and embodies the idea of culpability.[15] *Mens rea*, in this respect, means a desire or a will to cause a certain consequence as a result of the conduct of a person.[16] It reflect the offender's state of mind. What about robot crime? Can we treat autonomous intelligent machines' crimes under the existing criminal law principles?

As *Calo* points out,

> "*humans are, or are not, like robots, is a critical distinction that informs the legal issue before the court*".[17]

In this respect, the relevant question is: *Are AI systems capable of meeting the requirements of both factual and mental elements?* Although both elements should be present to establish criminal liability,[18] the greatest challenge remains the establishment of the *mens rea* element. How do we demonstrate the intentions of a non-human? This is particularly difficult due to AI systems' lack of consciousness.

---

[14] Hallevy, G. (2013) *When Robots Kill: Artificial Intelligence under Criminal Law*. Boston: Northeastern University Press, p. 85.

[15] Ibid.

[16] Gaur, K.D. (2003) Principles of Criminal Liability. In: K.D. Gaur (ed.). *Criminal Law and Criminology*. New Delhi: Deep and Deep Publications, p. 24.

[17] Calo, R. (2016) Op. cit., p. 25.

[18] Hallevy, G. (2013) Op. cit., p. 85.

Generally, self-consciousness represents a person's ability to think and make moral judgments, for instance judge good and evil. Thus, from an ethical and legal perspective, holding AI systems criminally accountable for their actions proves to be inappropriate, because they are not aware of the consequences of their actions. Such offenders are considered absent of adequate culpability (*nullum crimen sine culpa*).[19] However, from a technical perspective, researchers believe that robots already have a rudimentary consciousness, referring to robots' abilities to accumulate knowledge and make judgments accordingly,[20] and may soon become human-like entities, or even exceed human intelligence.[21] In light of that, *Hallevy* suggests that a new subject be added to criminal law, in addition to humans and corporations, to which he refers as *machina sapiens criminalis*.[22]

2.1. LEGAL PERSONHOOD:
CAN AI SYSTEMS BE SUBJECTS OF LAW?

Many researchers are looking for regulatory means to accommodate the status of intelligent agents in the legal system and in society. Perhaps encouraged by the case of two chimpanzees (*Hercules* and *Leo*) who were granted the status of "legal persons" by the *Manhattan Supreme Court* in 2015,[23] researchers have started to look at the possibility of granting "personhood" to AI systems. As one might expect, scholars share differing perspectives. It primarily depends on whether they accept the robot as a legal subject, i.e. *"metaphors or similes that support a particular verdict"*[24] or see it just as an object, i.e. *"artifacts in the world that have legal disputes."*[25]

---

[19]   Hallevy, G. (2015) *Liability for Crimes Involving Artificial Intelligence Systems*. Springer International Publishing, p. 33.

[20]   See, Retto, J. (2017) *Sophia, First Citizen Robot of the World*. [online] Available from: https://www.researchgate.net/publication/321319964_SOPHIA_FIRST_CITIZEN_ROBOT_O F_THE_WORLD [Accessed 29 November 2018] .

[21]   Radutniy, O.E. (2017) Op. cit., p. 136.

[22]   Hallevy, G. (2013) Op. cit., p. 21 (arguing that given the human-like actions of autonomous systems, the current criminal law is adequate to address AI technology issues. *Hallevy* continues his argument by stating that as long as technology is about to advance into creating what he calls "machina sapiens" – an AI system that imitates the human mind – current criminal law would be even more adequate to address criminal liability issues of AI entities, since the human mind is already the subject of the current criminal law).

[23]   Bekoff, M. (2015) *Judge Recognizes Two Chimpanzees as Legal Persons: A First Two Chimpanzees, Hercules and Leo, Are Determined to Be Persons in NY Court*. [online] Available from: www.psychologytoday.com/us/blog/animal-emotions/201504/judge-recognizes-two-chimpanzees-legal-persons-first [Accessed 29 November 2018].

[24]   Calo, R. (2016) Op. cit., p. 44.

[25]   Ibid.

Some legal scholars argue that robots are objects created by people and, as such, they are perceived as items or property, which by default cannot possess any legal rights.[26] They are merely regarded as tools in the hands of their producers or owners. This perspective reflects the views of the American writer *Isaac Asimov*, who distinguishes robots as property in the disposal of their human creators.[27]

In contrast, others note that AI systems are significantly different from other objects due to their active intervention in human relations. *Cerka et al.*, for example, point out that AI systems interact with other subjects of law, thus it is imperative to recognize them as legal entities in order to be able to clearly define subjects' rights, interests, and their obligations.[28] Such rights may prove difficult to enforce and thus call for establishment of legal provisions that accommodate the relationship between individuals and AI systems. An interesting case, for instance, was represented by a mock trial at the *International Bar Association* conference in San Francisco on 16 September 2013, where attorney *Dr. Martine Rothblatt* addressed the legality of a corporation unplugging/shutting down an intelligent computer, arguing that the computer had the right to maintain an existence.[29] In the absence of legislation, the court could not rule whether the AI system had the right not to be destroyed.

At the international level, robotic rights have already gained importance. In the autumn of 2017, Saudi Arabia granted citizenship to a humanoid robot named *Sophia*.[30] *Sophia*, a robot that mimicked human expressions and

---

[26]   See, for example: Radutniy, O.E. (2017) Op. cit., p. 136.

[27]   *Asimov* explores the "three law of robotics", which are as follows: *"(1) a robot may not injure a human being or, through inaction, allow a human being to come to harm; (2) a robot must obey orders given it by human beings except where such orders would conflict with the First Law; and (3) a robot must protect its own existence as long as such protection does not conflict with the First or Second Law"*. See: Houvast, F. Timmerman, R. and Zwaan, Y. (2018) Exploring the Legal Rights & Obligations of Robots: A Legal Book Review of I, Robot by Isaac Asimov reviewed in *Law Literature Project, Utrecht University*, pp. 1–8.

[28]   Having regard that AI systems cannot express their will as humans do, the authors suggest that by using legal analogy, AI entities should be viewed as artificial subjects of law, similar to the corporate personality. See: Čerka, P., Grigienė, J. and Sirbikytė, G. (2017) Is It Possible to Grant Legal Personality to Artificial Intelligence Software Systems?. *Computer Law and Security Review*, 33 (5), pp. 685–699.

[29]   Rothblatt, M. and Angelica, A.D. (2003) *Bio-Cyber-Ethics: Should We Stop a Company from Unplugging an Intelligent Computer?*. [blog entry] 28 September. Available from: http://www.kurzweilai.net/biocyberethics-should-we-stop-a-company-from-unplugging-an-intelligent-computer [Accessed 24 October 2018].

[30]   Wootson, C. (2017) *Saudi Arabia, Which Denies Women Equal Rights, Makes a Robot a Citizen*. [online] Available from: https://www.ndtv.com/world-news/saudi-arabia-which-denies-women-equal-rights-makes-a-robot-a-citizen-1768666 [Accessed 20 January 2019].

could learn human behaviours by interacting with people, became the first robot in the world to received citizenship.

The leading country in robotics, Japan, also seems to have settled the question of robot rights. On 7 November 2010, a therapeutic robot named *Paro*, received its own koseki, a special residency permit, which in Japanese society conflates legal rights such as family, nationality, and citizenship.[31]

In its turn, the *European Parliament* adopted a resolution entitled *Civil Law on Robotics*. This resolution envisages granting robots a specific legal status as electronic persons, responsible for any damage resulting from their autonomous decisions.[32] The assumption is that such a legal status would allocate liability for the damages caused by *"the most sophisticated autonomous robots"*, whose decisions are made independently, without human intervention.[33]

The *European Union's* oversight on granting autonomous systems legal personhood could be a starting point for allocating rights and responsibilities to AI systems; however, this insight is not accepted by all scholars. In March 2018, *Nathalie Nevejans*, a lecturer in law at the *University of Artois*, addressed this issue with an open letter to the *European Commission*, objecting to the legal status of robots.[34] She claims that a legal status for robots would collide with human rights, such as: the right to dignity, the right to its integrity, the right to remuneration, or the right to citizenship.[35]

It is fair to suggest that at some point in the evolution of AI, robots might become capable of generating their own goals and intentions, thus they may qualify as subjects of current criminal liability law. The next question is: How does the court punish an AI robot? Can they experience blame?

These issues remain without a clear answer. The traditional criminal justice system is designed by humans for humans, making punishment

---

[31] Robertson, J. (2014) Human Rights vs. Robot Rights: Forecasts from Japan. *Critical Asian Studies*, 46 (4), pp. 571–598.

[32] Civil Law Rules on Robotics, European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 16 February. Available from: http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2017-0051&language=EN&ring=A8-2017-0005 [Accessed 03 January 2019].

[33] Paragraph 59(f) of the Resolution.

[34] Robotics Open Letter. (2017) *Open Letter to the European Commission Artificial Intelligence and Robotics*. [online] Available from: http://www.robotics-openletter.eu/ [Accessed 20 January 2019].

[35] Ibid.

seem out of place in the discussion of robot field. Criminal law seeks justice by punishing offenders. Its key purpose is to establish security and order in society. In this respect, the role of criminal law is to correct wrongdoers, and to reintegrate them into society. Clearly, this may not be the case with robots, as they are different from humans in many aspects. Many human attributes could not apply to robots, such as: fear, pain, freedom of movement, etc. In this respect, robot imprisonment would be pointless. Therefore, the question of robot "personhood" remains an open debate with inconclusive results.

3. STRICT LIABILITY: AI AS A TOOL

In legal research, many scholars share the opinion that the legal responsibility for the actions of AI systems falls on the individual who has produced or owns the robot. For instance, *Asaro* explains that robots qualify as built systems where the schemes are chosen by designers, so the designers and programmers of the learning methods would be responsible for the machine's actions.[36] *Hallevy* expands on this idea. He introduces the *Perpetration-by-Another Liability Model*.[37] According to this model, AI entities are considered innocent agents – machines with no human attributes – and therefore an AI system could not be regarded as a perpetrator of an offence. Pursuant to this model, the real perpetrator is the programmer of the robot, who designs a software with certain instructions leading the robot to commit certain offences. In this scenario, although the conduct (*actus reus*) falls upon the AI, the intent (*mens rea*) is determined by the programmer's mental state.[38] The user of AI can also be considered as the perpetrator-by-another. *Hallevy* explains that AI is completely dependent on its user (end-user) who uses it with the intent to commit offences.[39]

It is important to note that this model suggests that AI is merely a tool (equivalent to a hammer) in the hands of the programmer or the user. It

---

[36] Asaro, P.M. (2006) What Should We Want From a Robot Ethic? *International Review of Information Ethics*, 6 (12), pp. 2–15. Available from: http://cybersophe.com/writing/Asaro IRIE.pdf [Accessed 20 January 2019].

[37] Hallevy, G. (2010) Op. cit., pp. 9–12.

[38] Ibid.

[39] Ibid.

lacks the *mens rea* element, hence the fault cannot be attributed to the AI itself.[40]

This perspective imposes non-fault liability and is similar to product liability regulated by the *Directive 85/374/EEC*.[41] According to this Directive, strict liability is imposed as a result of a manufacturing defect.[42] If the defective product causes an injury, producers are held strictly liable for that injury.[43]

Relying on the abovementioned analysis, AI systems are regarded as tools in the hands of the sellers, distributors, and users, and liability is imposed, not because of their wrong acts, but due to their specific relationship with the AI systems. The behaviour of an AI system is directly linked with the natural person on whose behalf it acts, therefore that person is held accountable for any injury caused by the AI system, regardless of whether such conduct was intentional or planned. This view reflects the concept of vicarious liability, according to which someone is held liable for the wrongs of others, such as, for example, employers who are vicariously liable for the wrongdoings of their employees.[44] Summarizing the idea of robot-as-tools based on the concept of vicarious liability, AI systems are regarded as agents acting on behalf of several individuals that could be considered as perpetrators, such as: the producer, the programmer, the designer, the seller, the distributor, the user of the robot, etc.

Yet, by comparing the parent-child relationship, *Sparrow* implies that product liability could not apply to programmers and robots, as it would be analogous to

---

[40] Freitas, P.M., Andrade, F. and Novais, P. (2014) Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible. In: Pompeu Casanovas, Ugo Pagallo, Monica Palmirani and Giovanni Sartor (eds.). *AI Approaches to the Complexity of Legal Systems*. Berlin: Springer, p. 150.

[41] 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. *Official Journal of the European Union* (L 210), 25 July. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31985L0374 [Accessed 23 January 2019].

[42] Article 1 states: *"The producer shall be liable for damage caused by a defect in his product"*. Ibid.

[43] Ibid.

[44] For an overview of vicarious liability, see: Lederman, E. (2000) Models for Imposing Corporate Criminal Liability: From Adaptation and Imitation Toward Aggregation and the Search For Self-Identity. *Buffalo Criminal Law Review*, 4 (1), pp. 641–708.

> *"holding parents responsible for the actions of their children once they have left their care".*[45]

Perhaps this reasoning should be given importance in the liability chain, due to AI systems' ability to make independent decisions and change behaviour through machine learning, making it very difficult to determine the cause-effect link between AI and the responsible participant. In this respect, strict liability might be a too heavy a burden on the producers or owners of the robot. Besides that, substantially different ways to resolve liability issues lead to potential legal errors.[46] Courts are likely to adopt different solutions, leading to widespread disagreement in the general public and among manufacturers trying to assess their potential liability in the national market.[47] Furthermore, an excessive blameworthiness could obstruct the progress of technology development and all its benefits would be lost.[48]

The practical solution lies in devising mechanisms of accountability that promote innovation and development of advanced autonomous artificial agents that benefit society, while effectively managing the risks and their harmful actions. In addressing the need to incentivise safety without ruining robot supply in the market, *Calo* argues in favour of providing protection to manufacturers and distributers of open robotics, similar to the immunity of gun manufacturers.[49]

3.1. NEGLIGENCE-BASED LIABILITY

Negligence imposes liability only if the defendant is at fault.[50] To simplify, liability in negligence in a criminal context arises if there is a reasonable

---

[45]   Sparrow, R. (2007) Killer Robots. *Journal of Applied Philosophy*, 24 (1), p. 62.

[46]   Geistfeld, M.A. (2017) A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation. *California Law Review*, 105, p. 1611.

[47]   Ibid.

[48]   Čerka et al. (2017) Op. cit., p. 689; See also: Gless, S. Silverman, E. and Weigend, T. (2016) If Robots Cause Harm, Who Is to Blame? Self-Driving Cars and Criminal Liability. *New Criminal Law Review*, 19 (3), p. 430.; see also Hubbard, P. (2014) 'Sophisticated Robots': Balancing Liability, Regulation, and Innovation. *Florida Law Review*, 66 (5), pp. 1803–1872 (explaining that *"All technology presents the challenge of balancing its costs against its benefits. First, the expanded potential liability of innovators could negatively affect their decision to develop, for example, an Unmanned Aircraft System (UAS). Second, the increased liability could also reduce the demand for a robotic vehicle like a UAS because purchasers and users would need to worry about potential greater liability for personal injury"*).

[49]   Calo, R. (2011) Open Robotics. *Maryland Law Review*, 70 (3), pp. 101–142. Available from: http://ssrn.com/abstract=1706293 [Accessed 25 January 2019].

[50]   Kelley, R. et al. (2010) Liability in Robotics: An International Perspective on Robots as Animals. *Advanced Robotics*, 24 (13), pp. 1861–1871.

duty of care. A negligent action is something a rational individual would not do. For example, texting on the phone while driving qualifies as negligence.

Negligence is the failure to use reasonable care to prevent harm to oneself or to others. A person can be negligent by acting or by failing to act. A person is negligent if he or she does something that a reasonably careful person would not do in the same situation or fails to do something that a reasonably careful person would do in the same situation.[51]

It seems logical that a person who has suffered a harm by an AI system files a suit against the agent that failed to provide due care, i.e. to do something that is morally and legally required. Other researchers share the same opinion. *Roff*, for instance, points out that the responsibility and liability lie with the software programmers. She discusses that the burden of legal responsibility falls upon programmers or the manufacturers, who, having the status of the creators of the machine, are legally responsible for providing warnings and taking necessary measures to avoid any risk.[52] In accordance with this view, *Hallevy* argues that the programmer or the user of the robot is liable when he or she has failed to foresee the offence committed by the robot and failed to prevent it.[53] In *Hallevy's* research, this is called the *Natural-Probable-Consequence liability model*.[54] According to this model, there is a link between the AI system's offence and manufacturer's or user's action. Even though the manufacturer or the user had no intent to commit an offence, if there is evidence that they should have taken corrective measures to prevent the happening, they are held accountable.[55] This model is based on the duty of reasonable programmers and manufacturers to foresee potential wrongdoings of the robots, i.e. to ensure that their products have no design flaws and to avoid obstacles by warning consumers of hidden dangers.

Criminal liability based on negligence relies on the assumption that none of the aforementioned individuals intended to cause harm to another

---

[51] Judicial Council of California. (2017) *California Civil Jury Instruction 2017 Edition ("CACI") No.401*. [online] Available from: https://www.justia.com/trials-litigation/docs/caci/400/401/ [Accessed 20 December 2018].

[52] Roff, H.M. (2013) Killing in War: Responsibility, Liability and Lethal Autonomous Robots. In: Fritz Allhoff, Nicholas Evans and Adam Henschke (eds.). *Routledge Handbook of Ethics and War: Just War Theory in the 21st Century*. Routledge Press, p. 356.

[53] Hallevy, G. (2010) Op. cit., pp. 15, 16.

[54] Ibid, p. 9.

[55] Ibid, pp. 14, 15.

human being, thus the liability arises due to lack of reasonable care, i.e. a rational person should have acted to avoid damage. This resembles the relationship between owners and domesticated animals.[56] *Kelley at al.* note that

> *"robot owners should be held liable for negligence with respect to their robots, much like dog owners are held liable for negligence with respect to their dogs".*[57]

In this sense, manufacturers, sellers, or distributors are criminally liable when they have failed to warn about a foreseeable risk or they have failed to give adequate instructions to avoid harm.

However, it is very difficult to impose liability based on negligence. Most importantly, it is not clear how to determine the standards of care. Let us imagine that instructions to avoid a risk in a given situation are provided in due time. The next question is: Was that warning or instruction reasonable? Currently there are no international legal norms determining safety requirements of AI systems upon which plaintiffs can rely in a certain situation. The acquisition of autonomy and the self-learning capacity of AI systems makes it difficult to predict their behaviours and to determine possible risks and threats, therefore it is not an easy task for victims who have the burden of proof to allocate the "fault". For example, if someone is injured by an autonomous robot that is designed to learn and to interact with the environment, it may be hard to identify what went wrong and what caused the accident. Additionally, due to the large number of individuals involved in the process of manufacturing, it is unclear which agent among multiple component suppliers of hardware and software has the duty to avoid risks.

3.1.1. FORESEEABILITY

To impose criminal liability based on negligence, it is important to determine whether the manufacturer could anticipate the potential malfunctions of the AI system. Prediction theory is central to negligence.[58] Negligence arises when a reasonably prudent person "ought to have

---

[56] Kelley, R. et al. (2010) Op. cit., p. 1864 (explaining: *"We justify this negligence analysis by analogizing robots to domesticated animals, whose owners are as a general rule subject to negligence liability because such animals are generally predictable"*).

[57] Ibid.

[58] Karnow, C.E.A. (2013) Op. cit., p. 9.

known" that his actions would cause harm.[59] In traditional product liability, the manufacturer is responsible for the product if he could foresee potential problems or harms it may cause. The verb "to foresee" means "to have prescience of; to know in advance".[60]

Foreseeability

"is not to be measured by what is more probable than not, but includes whatever is likely enough in the setting of modern life that a reasonably thoughtful [person] would take account of it in guiding practical conduct."[61]

Early robotic machines, such as automated elevators, have been subject to liability claims based on the duty of the manufacturer to provide reasonable care, i.e. to undertake safety measures to prevent risk or harm. Plaintiffs have sued for injuries caused by lack of due care in the maintenance of these machines.[62] However, there are no easy guidelines to determine whether the risk or physical harm is "foreseeable" for sophisticated and autonomous robots.

To effectively address responsibility and blame, as noted by *Owen*, we must first determine the individual's ability to understand that the action might cause harm and that opting for another action would be a safer alternative, in accordance to individual's obligations to the community.[63] The responsibility of manufacturers, distributors, and users of AI systems thus depend upon their capacity to understand the behaviour patterns of AI systems, the causal possibilities of AI systems' actions, and expected results. When it comes to intelligent robots and their ability to learn from environmental data, designers and users may not have a feasible way to predict their behaviour, simply because machine learning techniques

---

[59] Ibid, p. 10.

[60] Webster's Unabridged Dictionary. (2001) 2nd ed.

[61] *Constance B. v. State of California.* (1986) 178 Cal. App. 3d 200,206. Available from: https://law.justia.com/cases/california/court-of-appeal/3d/178/200.html [Accessed 20 December 2018].

[62] See, e.g. *Estabrook v. J.C. Penney Co.* (1970) 464 P.2d 325 (In Banc.). Available from: https://www.courtlistener.com/opinion/1247130/estabrook-v-jc-penney-company/ [Accessed 20 December 2018] (The plaintiff sued for injuries incurred while he was playing on an escalator); *Brown v. Sears, Roebuck & Co.* (1987) 514 So. 2d 439. Available from: https://law.justia.com/cases/louisiana/supreme-court/1987/87-c-0726-0-1.html [Accessed 20 December 2018] (A twenty-two month old child caught his finger in the space between the moving treads and the left side panel of the elevator. The court ruled in favor of the plaintiff).

[63] Owen, D. (2009) Figuring Foreseeability. *Wake Forest Law Review*, 44, pp. 1277–1307.

allow AI systems to modify their functions and learn after they are deployed.[64]

The second factor in determining responsibility is evaluation of the actor's decision;[65] whether it was the right decision that would have prevented the risk. This is often left to the judges to decide. As stated earlier, foreseeability is an abstract and vague concept, and in many cases cannot be proven in court. Some argue that judges use the doctrine of foreseeability as a means to weed out cases they deem unworthy;[66] that is, the judges use "foreseeability" as a means of deciding whether the action was right or wrong, upon which they impose or deny negligence liability. This can lead to different judgments for similar cases.[67]

Having said that, concepts like "foreseeability" and "reasonable care" cannot be used as key elements to impose criminal liability on manufacturers and sellers, as they may simply not have the level of skill required to foresee the manner in which the harm will occur. Consequently, standards of reasonable care may be vague. In addition, any attempt to impose responsibility on such a basis could lead to infinite liability for creators of AI systems that could obstruct the economy and innovation of AI.

4. AI LIABILITY GAP:
A LOOK INTO CORPORATE CRIMINAL LIABILITY

So far, we have seen that imposing criminal liability for the harmful and erratic behaviours of robots often leads us to a vicious circle due to the fact that it is often very difficult to attribute responsibility to specific individuals. When addressing the AI accountability gap, there are always two parties in the centre of the debate: those who develop AI, i.e. big

---

[64] Asaro, P.M. (2016) The Liability Problem for Autonomous Artificial Agents. In: Bipin Indurkhya and Georgi Stojanov (eds.). *AAAI Symposium on Ethical and Moral Considerations in Non-Human Agents*, Stanford, 21–23 March. USA: Association for the Advancement of Artificial Intelligence, pp. 190–194. Available from: https://www.aaai.org/ocs/index.php/SSS/SSS16/paper/view/12699 [Accessed 29 December 2018].

[65] Ibid.

[66] See, e.g. Cardi, W.J. (2005) Purging Foreseeability: The New Vision of Duty and Judicial Power in the Proposed Restatement (Third) of Torts. *Vanderbilt Law Review*, 58 (3), p. 739. Available from: https://wakespace.lib.wfu.edu/handle/10339/58895 [Accessed 15 October 2018] (explaining *"In many courts the foreseeability lens seems to expand, contract or change focus at the will of the judge"*); See also: Owen, D. (2009) Op. cit., p. 1278 (arguing *"while foreseeability may be the fundamental moral glue of tort, it provides so little decisional guidance that scholars often revile it for being vague, vacuous, and indeterminate"*).

[67] Cardi, WJ. (2005) Op. cit., p. 740.

companies; and those who are directly impacted from such development, i.e. society.[68] Anticipating an exponential growth of technology, it seems that this gap is going to accelerate in the future. Importantly, this paper examines whether accountability should be imposed over institutions, i.e. corporations, rather than individuals.

The need for corporate liability cannot be understood without an understanding of the great dominance and influence of the private sector. There are two key factors that ensure corporations' great digital power and dominance: data collection and money.

With the advent of globalisation, multinational corporations[69] wield tremendous financial power. As noted by *Pasquale*, these dominant players control money and information.[70] Today a great amount of digital power is concentrated in the hands of just a few people. This is a threat both to democracy and to functioning markets.[71] At the core of their reputation,[72] search,[73] and finance[74] lies our data, which too often is stored and processed under secrecy.[75] The big tech companies, in particular the "Frightful 5"[76] (*Google*, *Facebook*, *Amazon*, *Apple*, and *Microsoft*) control and collect the data and information of billions of people,[77] which is often used to shape consumers' behaviour and maximize profit.[78] Indeed, as *Bryson and Theodorou* point out,

---

[68] See: Whittaker, M. et al. (2018) *AI Now Report*. [online] New York City: AI Now Institute. Available from: https://ainowinstitute.org/AI_Now_2018_Report.pdf [Accessed 19 December 2018].

[69] The term "multinational corporation" is defined as *"an economic entity, which owns (in whole or in part), controls and manages income generating assets in more than one country"*. For the detailed review, see: Muchlinski, P.T. (2007) *Multinational Enterprises and the Law*. Oxford: Oxford University Press, p. 12.

[70] Pasquale, F. (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, Massachusetts: Harvard University Press, pp. 4–6.

[71] Nemitz, P. (2018) Constitutional Democracy and Technology in the Age of Artificial Intelligence. [in press] Submitted to: *Philosophical Transactions of the Royal Society A*. Available from: https://ssrn.com/abstract=3234336 [Accessed 07 January 2019].

[72] Pasquale, F. (2015) Op. cit., p. 4.

[73] Ibid.

[74] Ibid.

[75] Ibid.

[76] See: Manjoo, F. (2016) Tech's 'Frightful 5' Will Dominate Digital Life for Foreseeable Future. *N.Y. Times*, 20 January. [online] Available from: https://www.nytimes.com/2016/01/21/technology/techs-frightful-5-will-dominate-digital-life-for-foreseeable-future.html [Accessed 10 January 2019].

[77] Chadwick, P. (2018) To Regulate AI We Need New Laws, Not Just a Code of Ethics. *The Guardian*, 28 October. [online] Available from: https://www.theguardian.com/commentisfree/2018/oct/28/regulate-ai-new-laws-code-of-ethics-technology-power [Accessed 20 January 2019].

[78] Nemitz, P. (2018) Op. cit.

> *"with great power (or even just money) does come great responsibility".*[79]

Consequently, the big question is: Should corporations be held accountable for criminal actions arising from the tools they deploy in the market?

The "deep pockets" of the corporations might give us the answer to this question. Organisational blameworthiness, in this regard, derives from the profit-driven expansion of mega corporations' commercial activities. It seems legitimate and necessary to allocate the threats by recurring accountability to those who collect the fruits of AI deployment. Otherwise, the struggle to impose liability for the harmful actions of AI systems would provide room for big corporations to expand their businesses, while potential harmful acts of intelligent machines would remain a constant threat to society.

As noted by a report published in 2018, the accountability gap in AI is favourable for the companies that create and deploy these technologies, at the expense of society's interest.[80] Hence, a limited and unregulated corporation liability may encourage corporations to take greater risks – socially harmful risks – in order to maximize their profit. *Anderson and Luchsinger* share the same concerns:

> *"Most AI tools are and will be in the hands of companies striving for power. Values and ethics are often not baked into digital systems making people's decision for them. These systems are globally networked and not easy to regulate or rein in".*[81]

The big tech lobby campaigns against the law[82] give credence to these concerns. It demonstrates their efforts to avoid responsibility at all cost. The internet giants are the only corporations in history that have managed

---

[79] Bryson, J. and Theodorou, A. (2019) How Society Can Maintain Human-Centric Artificial Intelligence. In: Marja Toivonen and Eveliina Saari (eds.). *Human Centered Digitalization and Services*. Singapore: Springer, pp. 305–323. Available from: http://www.cs.bath.ac.uk/~jjb/ftp /BrysonTheodorou19.pdf [Accessed 20 January 2019].

[80] Whittaker et al. (2018) Op. cit.

[81] Anderson, J. Rainie, L. and Luchsinger, A. (2018) *Artificial Intelligence and the Future of Humans*. [online] Washington, DC: Pew Research Center. Available from: http://www.pewinternet.org/2018/12/10/artificial-intelligence-and-the-future-of-humans/ [Accessed 23 January 2019].

[82] For the latest big tech lobby campaigns against new law, see: Meyer, D. (2017) *Inside the EPrivacy Regulation's Furious Lobbying War*. [online] Kansas City: IAAP. Available from: https://iapp.org/news/a/inside-the-eprivacy-regulations-furious-lobbying-war/   [Accessed 23 January 2019]; Fang, L. (2018) *Google and Facebook Are Quietly Fighting California's Privacy Rights Initiative, Emails Reveal*. [online] First Look Media. Available from: https://the intercept.com/2018/06/26/google-and-facebook-are-quietly-fighting-californias-privacy-rights-initiative-emails-reveal/ [Accessed 23 January 2019].

to maximize their profit at the top of the stock exchange, and at the same time maintain their output largely unregulated.[83] Therefore, to shape their practice in the common good, there is an urgent need for sector-specific regulation. Without a regulation that goes beyond ethical standards, principles like justice and accountability will remain conflict concepts between mega-companies and the common good of the general public. Ethical codes, without enforceable mechanisms, have little or no effect regarding AI's social implications and liability problems. It seems that some of these companies have realized the complexity of the problem and have come to the same conclusion. A representative of *Microsoft*, for instance, calls for establishment of a new field of law governing *"a growing pool of businesses involved"*.[84]

Following the great impact corporations have in the development of AI, coupled with the difficulty to assign individual criminal liability and moral responsibility for crimes committed by intelligent robots, the discussion about holding corporations, rather than individuals, responsible is both timely and topical. The nature of the debate should now move on from "whether" legal entities should have criminal liability, to "how" to properly address and regulate their responsibility in the field of AI.

4.1. CONSIDERING CORPORATE CRIMINAL LIABILITY

As an initial matter, corporate liability is not a new concept. The advances in communication technologies and the rise of consumer activism in the mid-twentieth century brought a public debate regarding the impact of large multinational enterprises towards society.[85] The rise of social responsibility movement caused a shift from the state-centred concerns of corporate regulations (focusing on problems such as taxation and corruption) towards people-centred approaches, mainly focusing on health and safety issues.[86]

Given the great impact corporations have had on the economy, particularly railroads, the need to impose criminal liability to large

---

[83]   Nemitz, P. (2018) Op. cit.

[84]   Bass, D. (2018) *Microsoft Says AI Advances Will Require New Laws, Regulations*. [online] New York City: Bloomberg L.P. Available from: https://www.bloomberg.com/news/articles/2018-01-18/microsoft-says-ai-advances-will-require-new-laws-regulations [Accessed 21 January 2019].

[85]   Zerk, J.A. (2006) *Multinationals and Corporate Social Responsibility: Limitations and Opportunities in International Law*. New York: Cambridge University Press, p. 21.

[86]   Ibid, p. 23.

corporations was acknowledged by courts in the late nineteenth century. The courts began to apply criminal sanctions to corporations for actions committed by individuals.[87] In the years to come, the traditional principle of *societas delinquere non potest*[88] gradually faded away. As a result, most national jurisdictions recognise corporations as legal persons, and have put in place domestic legislation that regulate the imposition of criminal liability of corporations.[89] From a legal standpoint, then, corporate liability could fill the gap of accountability issues regarding autonomous systems.

There are several theories widely adopted to impose criminal liability of corporations.[90] Generally, corporate liability is imposed indirectly, through the acts of their agents, such as, for instance, vicarious liability and *respondeat superior*.[91] This represents a nominalist approach under which corporation's liability is induced from the individual liability of its representatives.[92] In these cases, a plaintiff would need to prove that there is an agent or group of agents upon whose erratic conducts the company's liability can be established. The liability of a company, therefore, does not represent an independent wrongdoing, but arises due to its legal relationship with these individuals.[93] The search for the rogue actor and individual blameworthiness brings to light complications inherent in assessing fault. As previously stated,[94] in response to the argument that an intelligent machine's acts are unpredictable, it remains a challenge as far as establishing individual intent. Consequently, it would be a difficult task for the courts to establish blame and seek justice. That would lead to absolute impunity.

---

[87] Dragatsi, H. (2011) *Criminal Liability of Canadian Corporations for International Crimes*. Canada: Thomson Reuters, p. 147.

[88] Literally means "corporations cannot commit crimes" and postulates that legal entities do not bear moral and criminal responsibility. See: Stoitchkova, D. (2010) *Towards Corporate Liability in International Criminal Law*. Utrecht: Utrecht University, p. 7.

[89] Ibid.

[90] For an overview of corporate criminal liability doctrines, see: Suhariyanto, B. (2018) Corporate Criminal Liability Under the Reactive Corporate Fault to Achieve Good Corporate Governance in Indonesia. In: A. Raharjo and T. Sudrajat (eds.). *The 1st International Conference on Law, Governance and Social Justice*, Purwokerto, 25–26 September. Les Ulis: EDP Sciences. Available from: https://www.shs-conferences.org/articles/shsconf/pdf/2018/15/shsconf_icolgas2018_07009.pdf [Accessed 20 January 2019].

[91] See: Lederman, E. (2000) Op. cit., p. 651.

[92] Colvin, E. (1995) Corporate Personality and Criminal Liability. *Criminal Law Forum*, 6 (1), p. 2.

[93] Ibid.

[94] See section 2 of this paper.

A new concept of corporate liability, undergirded by a realistic view of corporations, suggests that corporations have individual personalities and intentions, which does not derive from the actions of its agents.[95] The realistic view on corporate criminal liability could provide normative support for law reformers. American literature in the past decade proposed a new model of criminal liability of legal bodies, which is also based on self--identity doctrine and reflects the idea of modern corporations.[96] According to this approach, liability of corporations is primary. The corporate fault is structured by considering a multitude of factors, directly questioning corporation's conduct, i.e. what it ought to have known and what it ought do have done in order to prevent harm.[97]

In support of the individual identity of corporations approach, it can be argued that a theoretical shift from derivative models, and a new concept of strict corporate liability based on the independent identity of the legal body could alleviate issues at hand. Considering the emerging concept of corporations,[98] and the unprecedented power in their hands, a focus towards organisational blameworthiness could be a powerful tool in allocating responsibility for the risks associated with intelligent machines. The exponential development of technology and AI calls for regulatory framework to address situations not anticipated before. Policymakers acknowledge that legal entities, who have managed to attain, to a large extent, trouble free, should be subject to criminal proceedings.[99]

In the absence of a particular legislation, it is imperative to look for legal alternatives. For instance, corporate criminal liability could be developed

---

[95] See, e.g. Colvin, E. (1995) Op. cit. (*"Corporations can act and be at fault in ways that are different from the ways in which their members can act and be at fault"*.); Lawrence, F. (2000) In Defense of Corporate Criminal Liability. *Harvard Journal of Law & Public Policy*, 23 (3), p. 833 (discussing that corporations *"have independent identities in the community, based upon attributes-identifiable personae and a capacity to express moral judgments-that substantively distinguish them from their owners, managers and employees"*.); Bucy, P.H. (1991) Corporate Ethos: A Standard for Imposing Corporate Criminal Liability. *Minnesota Law Review*, 75, p. 1182 (proposing a corporate ethos doctrine based on the idea of separate corporate intent).

[96] Lederman, E. (2000) Op. cit., p. 678.

[97] Ibid.

[98] The complex form of modern organisations, their decentralised units, etc. often prevent identifying personal liability of individual agents; therefore, the idea that the corporations have an individual personality that does not derive from its representatives has gained momentum in recent decades. See: Sperino, S. (2010) A Modern Theory of Direct Corporate Liability for Title VII. *Faculty Articles and Other Publications Paper 230*. Available from: https://scholarship.law.uc.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&https redir=1&article=1233&context=fac_pubs [Accessed 21 January 2019].

[99] Lederman, E. (2000) Op. cit., p. 644.

through a correlate stream of public welfare doctrine for strict-liability offences.[100] This theory serves two purposes: First, it represents a doctrine that was designed as a necessity to promote social order at a time when the industrial revolution brought many new threats to society.[101] The public welfare rules were intended to control particular industries and activities that affected public health, safety or welfare.[102] Similarly, the digital revolution has brought so many threats unknown to humans before. Taking into consideration the great socio-economic impact of mega-corporations, the welfare offence doctrine could pave the way for charging corporations with criminal offences for the harmful acts of AI. Second, this doctrine omits the criteria of blameworthiness; a penalty can be imposed regardless of the actor's intent, so the plaintiff does not have to prove that the defendant acted purposely.[103] Intent or blameworthiness is replaced by the assumption of the risk that the actor bears when engaging in a certain activity.[104]

Assumption is particularly important, as it may resolve many core problems facing the liability regime of autonomous systems. A particular death caused by an intelligent robot, for instance, should alarm AI developers with respect to the risks associated with the tools they deploy in the market. Their failure to adopt necessary precautionary and corrective measures should be subject to criminal liability, provided that they should have known or at least should have assumed that their conduct may seriously threaten the community's safety. Responsibility, hence, arises as a burden to an entity,

> *"otherwise innocent, but standing in responsible relation to a public danger"*.[105]

Examples given by *Meyer* illustrate my suggestion:

A defendant charged with criminal possession of an unregistered firearm may be convicted even if he mistakenly thought the firearm was

---

[100] Strict liability offences are those where liability can be imposed without a *mens rea* of an actor. See section 2 of this paper.

[101] Carpenter, C.L. (2003) On Statutory Rape, Strict Liability, and the Public Welfare Offense Model. *American University Law review*, 53 (2), pp. 313–391.

[102] See: *Morissette v. United States.* (1952) 342 U.S. 246. Available from: https://supreme.justia.com/cases/federal/us/342/246/ [Accessed 28 January 2019].

[103] Reitz, M.J. (2013) *Strict Liability and Public Welfare Offenses*. [online] Midland: Mackinac Center. Available from: https://www.mackinac.org/19579 [Accessed 21 January 2019].

[104] Carpenter, C.L. (2003) Op. cit., p. 320.

[105] *Morissette v. United States.* (1952) Op. cit.

registered as required; An immigrant alien may be criminally convicted for unlawfully re-entering the United States even if she believed that she had proper government approval to return.[106]

In 1994, for instance, the *Supreme Court* in *Staples v. United States*[107] imposed strict liability to a defendant possessing an unregistered "firearm" that had been modified into an automatic weapon.[108] Although the defendant claimed he was unaware of the factual situation and was ignorant of any automatic capability of the firearm, the Court ruled that

> *"as long as a defendant knows that he is dealing with a dangerous device of a character that places him in responsible relation to a public danger, he should be alerted to the probability of strict regulation".*[109]

Is a public welfare offence doctrine the solution then? One could argue that it depends on whether you consider that a serious injury or a death caused by an intelligent robot is not only a harmful event for society, but also a probable consequence of a blameworthy negligence, which ought to be identified and repressed.

4.2. QUESTIONS AND DOUBTS

Commenting on corporate criminal responsibility, *Chengeta* has observed that corporate responsibility may bring jurisdictional and cost challenges to victims.[110] Victims would be faced with the burden of bringing cases against corporations registered in foreign countries, thereby encountering enormous financial costs.[111]

Additionally, it is beyond dispute that an excessive burden of criminal liability on corporations would impede innovation. Making corporations fully liable means that they should not invest in advancing AI. Does society really want this? Perhaps to some extent the general public needs to adapt to technological advancement. That is to say, probably there should be some

---

[106] Meyer, J.A. (2007) Authentically Innocent: Juries and Federal Regulatory Crimes. *Hastings Law Journal*, 59 (1), p. 138.

[107] *Staples v. United States.* (1994) 511 U.S. 600. Available from: https://supreme.justia.com/cases/federal/us/511/600/ [Accessed 28 January 2019].

[108] Ibid. at 602–04 ("Automatic" refers to a weapon that fires repeatedly with a single pull of the trigger).

[109] Ibid.

[110] Chengeta, T. (2016) Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law. *Denver Journal of International Law and Policy*, 45 (1), p. 4.

[111] Ibid.

safety issues that should be socially tolerated, especially when it comes to robots dedicated to social benefits. Some authors[112] postulate that we should look for an intermediate solution, by adjusting legal rules and by accepting a margin of tolerance of certain errors in the designing and programming of robots.[113]

Tolerance for robot malfunctions, though, should not play a central role in debates concerning the existence of corporate criminal liability. A variety of potential, harmful consequences will also occur if there is no one to be blamed for the dangers associated with the use of autonomous systems. Therefore, the challenge of balancing technological development and its benefits in relation to its potential threats remains a future challenge.

5. CONCLUSION

The successful development of the robot market has posed challenging liability issues that need to be addressed by regulators and policymakers. There is no one-size-fits-all solution. In the absence of an international regulatory framework, finding the person that should be held criminally liable for the mishaps of fully autonomous systems is not an easy task, especially due to the complex and sophisticated nature of the machines. The more autonomous the system, the greater the challenge to establish effective rules governing liability for harmful actions.

While some authors suggest that current criminal law is a plausible possibility to cope with AI,[114] this paper shows that liability questions[115] of AI systems push traditional criminal law to its limit. Old definitions are not suitable to the modern era.

In terms of thinking of robots as subject to the law, one could argue that it is not a straightforward solution, due to the uncertainty of many questions facing robot ethics. While giving civic status to AI systems would regulate the industry to some extent, many other problems, such as intention and culpability of AI systems, remain to be addressed.

Similarly, product liability is facing challenging questions that stem from the increasingly sophisticated nature of robots and the many individuals involved in the production and maintenance of AI systems. It cannot be

---

[112] Gless, S., Silverman, E. and Weigend, T. (2016) Op. cit., p. 17.
[113] Ibid.
[114] Hallevy, G. (2013) Op. cit., p. 29.
[115] Questions of risk, fault and punishment remain to be addressed.

conclusively established because methods to impose product liability are not clear.

We argue that there is a need for a shift from an individual-centred liability model to an organisation-centred liability model. We suggest that corporate criminal liability is the best solution whenever the responsibility of individual agents cannot be conclusively established. However, in the absence of a particular law in the AI field, this proposition is also fit for discussion.

The existing conflict clearly shows that AI systems and issues related to their criminal liability will remain a topic of discussion among academic for many years to come.

LIST OF REFERENCES

[1]   Anderson, J. Rainie, L. and Luchsinger, A. (2018) *Artificial Intelligence and the Future of Humans*. [online] Washington DC: Pew Research Center. Available from: http://www.pewinternet.org/2018/12/10/artificial-intelligence-and-the-future-of-humans/ [Accessed 23 January 2019].

[2]   Asaro, P.M. (2006) What Should We Want From a Robot Ethic?. *International Review of Information Ethics*, 6 (12), pp. 2–15. Available from: http://cybersophe.com/writing/ Asaro IRIE.pdf [Accessed 20 January 2019].

[3]   Asaro, P.M. (2016) The Liability Problem for Autonomous Artificial Agents. In: Bipin Indurkhya and Georgi Stojanov (eds.). *AAAI Symposium on Ethical and Moral Considerations in Non-Human Agents*, Stanford, 21–23 March. USA: Association for the Advancement of Artificial Intelligence, pp. 190–194. Available from: https://www.aaai.org/ocs/index.php/SSS/SSS16/paper/view/12699 [Accessed 29 December 2018].

[4]   Bass, D. (2018) *Microsoft Says AI Advances Will Require New Laws, Regulations*. [online] New York City: Bloomberg L.P. Available from: https://www.bloomberg.com/news/ articles/2018-01-18/microsoft-says-ai-advances-will-require-new-laws-regulations [Accessed 21 January 2019].

[5]   Bekoff, M. (2015) *Judge Recognizes Two Chimpanzees as Legal Persons: A First Two Chimpanzees, Hercules and Leo, Are Determined to Be Persons in NY Court.* [online] Available from: https://www.psychologytoday.com/us/blog/animal-emotions/201504/judge-recognizes-two-chimpanzees-legal-persons-first [Accessed 02 December 2018].

[6]   *Brown v. Sears, Roebuck & Co.* (1987) 514 So. 2d 439. Available from: https://law.justia. com/cases/louisiana/supreme-court/1987/87-c-0726-0-1.html [Accessed 20 December 2018].

[7]   Bryson, J. and Theodorou, A. (2019) How Society Can Maintain Human-Centric Artificial Intelligence. In: Marja Toivonen and Eveliina Saari (eds.). *Human Centered Digitalization and Services*. Singapore: Springer, pp. 305–323. Available from: http://www.cs.bath.ac.uk/ ~jjb/ftp/BrysonTheodorou19.pdf [Accessed 20 January 2019].

[8]   Bucy, P.H. (1991) Corporate Ethos: A Standard for Imposing Corporate Criminal Liability. *Minnesota Law Review*, 75.

[9]   Calo, R. (2011) Open Robotics. *Maryland Law Review*, 70 (3). Available from: http://ssrn.com/abstract=1706293 [Accessed 25 January 2019].

[10]   Calo, R. (2016) Robots in American Law. *University of Washington School of Law Research Paper No. 2016-042*. Available from: https://ssrn.com/abstract=2737598 [Accessed 22 June 2019].

[11]   Cardi, W.J. (2005) Purging Foreseeability: The New Vision of Duty and Judicial Power in the Proposed Restatement (Third) of Torts W. *Vanderbilt Law Review*, 58 (3). Available from: https://wakespace.lib.wfu.edu/handle/10339/58895 [Accessed 15 October 2018].

[12]   Carpenter, C.L. (2003) On Statutory Rape, Strict Liability, and the Public Welfare Offense Model. *American University Law review*, 53 (2), pp. 313–391.

[13]   Čerka, P., Grigienė, J. and Sirbikytė, G. (2017) Is It Possible to Grant Legal Personality to Artificial Intelligence Software Systems?. *Computer Law and Security Review*, 33 (5), pp. 685–699.

[14]   Chadwick, P. (2018) To Regulate AI We Need New Laws, Not Just a Code of Ethics. *The Guardian*, 28 October. [online] Available from: https://www.theguardian.com/ commentisfree/2018/oct/28/regulate-ai-new-laws-code-of-ethics-technology-power [Accessed 20 January 2019].

[15]   Chengeta, T. (2016) Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law. *Denver Journal of International Law and Policy*, 45 (1).

[16]   Civil Law Rules on Robotics, European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 16 February. Available from: http://www.europarl.europa.eu/sides/getDoc.do?type=TA& reference=P8-TA-2017-0051&language=EN&ring=A8-2017-0005 [Accessed 03 January 2019].

[17]   Colvin, E. (1995) Corporate Personality and Criminal Liability. *Criminal Law Forum*, 6 (1).

[18]   *Constance B. v. State of California.* (1986) 178 Cal. App. 3d 200, 206. Available from: https://law.justia.com/cases/california/court-of-appeal/3d/178/200.html
[Accessed 20 December 2018].

[19]   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. *Official Journal of the European Union* (L 210), 25 July. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31985L0374
[Accessed 23 January 2019].

[20]   Deep Knowledge Analytics. (2018) *Artificial Intelligence Industry in the UK.* [online] London: DKA. Available from: https://www.dka.global/ai-in-uk-report
[Accessed 12 May 2019].

[21]   Dragatsi, H. (2011) *Criminal Liability of Canadian Corporations for International Crimes*. Canada: Thomson Reuters.

[22]   *Estabrook v. J.C. Penney Co.* (1970) 464 P.2d 325 (In Banc.). Available from: https://www.courtlistener.com/opinion/1247130/estabrook-v-jc-penney-company/
[Accessed 20 December 2018].

[23]   Fang, L. (2018) *Google and Facebook Are Quietly Fighting California's Privacy Rights Initiative, Emails Reveal*. [online] First Look Media. Available from: https://theintercept.com/2018/06/26/google-and-facebook-are-quietly-fighting-californias-privacy-rights-initiative-emails-reveal/ [Accessed 23 January 2019].

[24]   Freitas, P.M., Andrade, F. and Novais, P. (2014) Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible. In: Pompeu Casanovas, Ugo Pagallo, Monica Palmirani and Giovanni Sartor (eds.). *AI Approaches to the Complexity of Legal Systems*. Berlin: Springer.

[25]   Gaur, K.D. (2003) Principles of Criminal Liability. In: K.D. Gaur (ed.). *Criminal Law and Criminology*. New Delhi: Deep and Deep Publications.

[26]   Geistfeld, M.A. (2017) A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation. *California Law Review*, 105.

[27]   Gless, S. Silverman, E. and Weigend, T. (2016) If Robots Cause Harm, Who Is to Blame? Self-Driving Cars and Criminal Liability. *New Criminal Law Review*, 19 (3).

[28]   Grace, K. et al. (2018) Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts. *Journal of Artificial Intelligence Research*, 62, pp. 729–754.

[29]   Hallevy, G. (2010) 'I, Robot – I, Criminal' – When Science Fiction Becomes Reality: Legal Liability of AI Robots Committing Criminal Offenses. *Syracuse Science and Technology Law Reporter*, 22 (spring).

[30]   Hallevy, G. (2013) *When Robots Kill: Artificial Intelligence under Criminal Law*. Boston: Northeastern University Press.

[31]   Hallevy, G. (2015) *Liability for Crimes Involving Artificial Intelligence Systems*. Springer International Publishing.

[32]   Houvast, F. Timmerman, R. and Zwaan, Y. (2018) Exploring the Legal Rights & Obligations of Robots: A Legal Book Review of I, Robot by Isaac Asimov reviewed in *Law & Literature Project, Utrecht University*.

[33]   Hubbard, P. (2014) 'Sophisticated Robots': Balancing Liability, Regulation, and Innovation. *Florida Law Review*, 66 (5), pp. 1803–1872.

[34]   Judicial Council of California. (2017) *California Civil Jury Instruction 2017 Edition ("CACI") No. 401*. [online] Available from: https://www.justia.com/trials-litigation/docs/caci/400/401/ [Accessed 20 December 2018].

[35]   Karnow, C.E.A. (1996) Liability for Distributed Artificial Intelligences. *Berkeley Technology Law Journal*, 11 (1), pp. 147–204. Available from: http://scholarship.law.berkeley.edu/btlj/vol11/iss1/3 [Accessed 29 November 2018].

[36]   Karnow, C.E.A. (2013) The Application of Traditional Tort Theory to Embodied Machine Intelligence. In: Ryan Calo, Michael Froomkin, Ian Kerr (eds.). *Robot Law*. Cheltenham: Edward Elgar Publishing Ltd, pp. 51–77.

[37]   Kelley, R. et al. (2010) Liability in Robotics: An International Perspective on Robots as Animals. *Advanced Robotics*, 24 (13), pp. 1861–1871.

[38]   Lawrence, F. (2000) In Defense of Corporate Criminal Liability. *Harvard Journal of Law & Public Policy*, 23 (3).

[39]   Lederman, E. (2000) Models for Imposing Corporate Criminal Liability: From Adaptation and Imitation Toward Aggregation and the Search for Self-Identity. *Buffalo Criminal Law Review*, 4 (1), pp. 641–708.

[40]   Manjoo, F. (2016) Tech's 'Frightful 5' Will Dominate Digital Life for Foreseeable Future. *N.Y. Times*, 20 January. [online] Available from: https://www.nytimes.com/2016/01/21/technology/techs-frightful-5-will-dominate-digital-life-for-foreseeable-future.html [Accessed 10 January 2019].

[41]   Meyer, D. (2017) *Inside the EPrivacy Regulation's Furious Lobbying War*. [online] Kansas City: IAAP. Available from: https://iapp.org/news/a/inside-the-eprivacy-regulations-urious-lobbying-war/ [Accessed 23 January 2019].

[42]   Meyer, J.A. (2007) Authentically Innocent: Juries and Federal Regulatory Crimes. *Hastings Law Journal*, 59 (1).

[43]   *Morissette v. United States*. (1952) 342 U.S. 246. Available from: https://supreme.justia. com/cases/federal/us/342/246/ [Accessed 28 January 2019].

[44]   Muchlinski, P.T. (2007) *Multinational Enterprises and the Law*. Oxford: Oxford University Press.

[45]   Murphy, M. (2017) Stephen Hawking: AI Could Be Best – or Worst – Thing in Human History. [online] New York City: MarketWatch. Available from: https://www.market watch.com/story/stephen-hawking-ai-could-be-best-or-worst-thing-in-human-history-2017-11-06 [Accessed 29 October 2018].

[46]   Nemitz, P. (2018) Constitutional Democracy and Technology in the Age of Artificial Intelligence. [in press] Submitted to: *Philosophical Transactions of the Royal Society A*. Available from: https://ssrn.com/abstract=3234336 [Accessed 07 January 2019].

[47]   Osborne, S. (2017) Elon Musk Calls for Urgent Laws on Robot as They Will Soon Be Risk to Public. *Express*, 28 November. [online] Available from: https://www.express.co.uk/ news/science/885344/elon-musk-artificial-intelligence-robotics-regulation [Accessed 29 October 2018].

[48]   Owen, D. (2009) Figuring Foreseeability. *Wake Forest Law Review*, 44, pp. 1277–1307.

[49]   Palmer, A. (2010) Autonomous UAS: a partial solution to America's future airpower needs. [pre-print] Submitted to: *Air University in partial fulfillment of the graduation requirements*. Available from: https://apps.dtic.mil/dtic/tr/fulltext/u2/1018416.pdf [Accessed 29 November 2018].

[50]   Pasquale, F. (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, Massachusetts: Harvard University Press.

[51]   Pratt, G.A. (2015) Is a Cambrian Explosion Coming for Robotics?. *Journal of Economic Perspectives*, 29 (3).

[52]   Radutniy, O.E. (2017) Criminal Liability of the Artificial Intelligence. *Problems of Legality*, 138.

[53]   Reitz, M.J. (2013) *Strict Liability and Public Welfare Offenses*. [online] Midland: Mackinac Center. Available from: https://www.mackinac.org/19579 [Accessed 21 January 2019].

[54]   Retto, J. (2017) *Sophia, First Citizen Robot of the World*. [online] Available from: https://www.researchgate.net/publication/321319964_SOPHIA_FIRST_CITIZEN_ROBOT _OF_THE_WORLD [Accessed 29 November 2018] .

[55]   Robertson, J. (2014) Human Rights vs. Robot Rights: Forecasts from Japan. *Critical Asian Studies*, 46 (4), pp. 571–598.

[56]   Robotics Open Letter. (2017) *Open Letter to the European Commission Artificial Intelligence and Robotics*. [online] Available from: http://www.robotics-openletter.eu/ [Accessed 20 January 2019].

[57]   Roff, H.M. (2013) Killing in War: Responsibility, Liability and Lethal Autonomous Robots. In: Fritz Allhoff, Nicholas Evans and Adam Henschke (eds.). *Routledge Handbook of Ethics and War: Just War Theory in the 21st Century*. Routledge Press.

[58]   Rothblatt, M. and Angelica, A.D. (2003) *Bio-Cyber-Ethics: Should We Stop a Company from Unplugging an Intelligent Computer?*. [blog entry] 28 September. Available from: http://www.kurzweilai.net/biocyberethics-should-we-stop-a-company-from-unplugging-an-intelligent-computer [Accessed 24 October 2018].

[59]   Sparrow, R. (2007) Killer Robots. *Journal of Applied Philosophy*, 24 (1).

[60]   Sperino, S. (2010) A Modern Theory of Direct Corporate Liability for Title VII. *Faculty Articles and Other Publications Paper 230*. Available from: https://scholarship.law.uc.edu/ cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1233& context=fac_pubs [Accessed 21 January 2019].

[61]   *Staples v. United States*. (1994) 511 U.S. 600. Available from: https://supreme.justia.com/ cases/federal/us/511/600/ [Accessed 28 January 2019].

[62]   Stoitchkova, D. (2010) *Towards Corporate Liability in International Criminal Law*. Utrecht: Utrecht University.

[63]   Suhariyanto, B. (2018) Corporate Criminal Liability Under the Reactive Corporate Fault to Achieve Good Corporate Governance in Indonesia. In: A. Raharjo and T. Sudrajat (eds.). *The 1st International Conference on Law, Governance and Social Justice, Purwokerto*, 25–26 September. Les Ulis: EDP Sciences. Available from: https://www.shs-conferences. org/articles/shsconf/pdf/2018/15/shsconf_icolgas2018_07009.pdf [Accessed 20 January 2019].

[64]   Toon, J. (2017) *Swarms of Autonomous Aerial Vehicles Test New Dogfighting Skills*. [online] Atlanta: GeorgiaTech. Available from: http://www.rh.gatech.edu/news/590743/swarms-autonomous-aerial-vehicles-test-new-dogfighting-skills [Accessed 15 May 2019].

[65]   Vojislav, K. (2001) *Learning and Soft Computing: Support Vector Machines, Neutral Networks and Fuzzy Logic Models*. Cambridge: The MIT Press.

[66]   Whittaker, M. et al. (2018) *AI Now Report*. [online] New York City: AI Now Institute. Available from: https://ainowinstitute.org/AI_Now_2018_Report.pdf [Accessed 19 December 2018].

[67]   Willick, M. (1983) Artificial Intelligence: Some Legal Approaches and Implications. *AI Magazine*, 4 (2). Available from: https://aaai.org/ojs/index.php/aimagazine/article/view/392 [Accessed 29 November 2018].

[68]   Wootson, C. (2017) *Saudi Arabia, Which Denies Women Equal Rights, Makes a Robot a Citizen*. [online] Available from: https://www.ndtv.com/world-news/saudi-arabia-which-denies-women-equal-rights-makes-a-robot-a-citizen-1768666 [Accessed 20 January 2019].

[69]   Zerk, J.A. (2006) *Multinationals and Corporate Social Responsibility: Limitations and Opportunities in International Law*. New York: Cambridge University Press.