LIABILITY OF ROBOTS: A NEW ERA OF CRIME

Rupal Dutta[1] and Archisha Suraj[2]

INTRODUCTION

What only decades ago was the stuff of science fiction is today the reality of science: computers capable of solving problems by replicating human thought processes.[3]

Could the liability of any mishap rest with the robot itself? AI devices are becoming increasingly autonomous. They derive their autonomy from a highly complex set of algorithms and their ability to react to unpredictable environmental scenarios. Therefore, an argument can be made that it would be unfair to impose liability on either the user or the manufacturer as they can't truly predict the robot's actions. If the devices could be found liable, what would be the appropriate punishment? A European study has suggested, in a report on RoboLaw, that if we were to hold robots liable for their actions, they would need to be registered and equipped with sufficient assets to be able to pay any compensation to injured parties. Professor Gabriel Hallevy has come up with possible defence arguments for robots, for example, if it has been infected with a virus (i.e., robotic diminished responsibility). The developments in AI will therefore require an examination of the established principles of liability. The regime which is ultimately created to determine liability will need to ensure that the principles are fair and address the interests of users and manufacturers, whilst not harming innovation.[4]

There have also been serious questions raised as to who could be held accountable for the deaths caused by autonomous weapons in war (Human Rights Watch 2015).[5]

Hong Kong-based technology law specialist Francois Tung of Pinsent Masons, the law firm behind Out-Law.com, said the law should recognise manufacturers or users of self-learning machines as ultimately responsible for what those machines do."We need to distinguish moral

[1] Student, School of Law, UPES
[2] Student, School of Law, UPES
[3] Maruerite Gerstner, "Liability Issues with Artificial Intelligence Software" by Maruerite E. Gerstner Digitalcommons.law.scu.edu (2017).
[4] LawCareers.Net. (2019). *Is the law ready for artificial intelligence?* Available at: https://www.lawcareers.net/Information/CommercialQuestion/Olswang-LLP-Can-a-robot-be-responsible-when-it-makes-a-mistake-AI-and-the-law.
[5] Peter Asaro(2017), The Liability Problem for Autonomous Artificial Agents.

responsibility and legal responsibility," Tung said. "Legal rules on liability have to be expanded to take into account AI. While it may be a matter of philosophical debate whether AI has 'free will', and hence the ability to intentionally cause harm to others and to be held responsible for such action, I believe that humans can never relinquish oversight of computers. Therefore the manufacturer or user of machines should bear ultimate responsibility, after all people create those machines and program them to work."[6]

HISTORICAL BACKGROUND

In 1942, **Isaac Asimov**, an American author of sci-fi novels and a professor of Biochemistry, was the first to try to devise a set of rules applicable to robots in a hypothetical future.

Although these laws are derived from fictional narrative, they are regarded by many as the most coherent attempt so far to establish a framework of laws in which smart independent robots can operate. These laws state that:

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 the orders given to it by human beings except where such orders would conflict with the First Law;

3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.[7]

Another law, added later by Asimov, states that ‘a robot may not harm humanity, or, by inaction, allow humanity to come to harm’.

The European Parliament has taken Asimov’s Laws as a foundation for formulating arguments regarding the regulatory framework for robotics. This reasoning behind this is that despite it being a work of fiction, until the moment when robots become or are made self-aware and an appropriate legislative framework is in place, Asimov’s Laws must be regarded as being directed

[6] Legal liability rules need to be expanded to account for artificial intelligence, says expert, Out-law.com (2017).
[7] Asimov, Isaac (1950). "Runaround". I, Robot (The Isaac Asimov Collection ed.). New York City: Doubleday. p. 40

at the designers, producers and operators of robots, since those laws cannot be converted into machine code.

The lacunae in the current legal framework is that, robots cannot be held liable per se for acts or omissions that cause damage to third parties. The scope of the existing rules on liability covers cases where the cause of the robot's act or omission can be traced back to a specific human agent such as the manufacturer, the owner or the user, and where that agent could have foreseen and avoided the robot's harmful behavior[8].

The draft report also proposes that a mandatory insurance scheme could be a possible solution to the complexity of allocating responsibility for damage caused by increasingly autonomous robots. However, unlike the insurance system for road traffic, where the insurance covers human acts and technical failures, an insurance system for robotics could be based on the manufacturer's perceived responsibility for the autonomous robots it produces.[9]

CRIMINAL LIABILITY

Three models of AI entity criminal liability: the perpetration-by-another liability model, the natural-probable-consequence liability model and the direct liability model, exists[10]. These three models might be applied separately, but in many situations, a coordinated combination of them (all or some of them) is required in order to complete the legal structure of criminal liability.[11]

A. The Perpetration-via-Another Liability Model: Artificial Intelligence Entities as Innocent Agents

This first model regards the AI entity as deprived of any human attributes making the assumption that AI entity is an innocent agent[12]. Accordingly, due to that legal viewpoint, a machine is a machine, and is never human. However, it would be erroneous to disregard an

[8] GonencGurkaynak, IlayYilmaz, GunesHaksever; Stifling artificial intelligence: Human perils computer law & security review 32 (2 0 1 6) p. 749 –758

[9]Policy, I. (2017). *Robots and liability issues: the future regulatory framework | Inline Policy*. [online] Inlinepolicy.com. Available at: https://www.inlinepolicy.com/ [Accessed 24 Jan. 2019].

[10]Hallevy, G. (2010). The Criminal Liability of Artificial Intelligence Entities. *SSRN Electronic Journal*, pp.179-183.

[11] BSN UK ltd. VsJanardan Mohandas RajanPillai [1996] 86 Com Cases 371 (Bom)s

[12]Hallevy, G. (2010). The Criminal Liability of Artificial Intelligence Entities. *SSRN Electronic Journal*, pp.179-183.

CC BY-NC 4.0

AI entity's capabilities. In consonance with the assumption of this model, these capabilities are insufficient to deem the AI entity a perpetrator of an offense. These capabilities are considered at par with capabilities of a mentally limited person, such as a child, a person who is mentally incompetent, or one who lacks a criminal state of mind[13].

Legally, when an offense is committed by an innocent agent (a child, a person who is mentally incompetent, or one who lacks a criminal state of mind to commit an offense) that person is criminally liable as a perpetrator-via-another. In such cases, the whole liability lies on the shoulder of the party orchestrating the offense (the perpetrator-via-another) who is deemed the real perpetrator as a principal in the first degree and is held accountable for the conduct of the innocent agent. The intermediary is regarded as a mere instrument, albeit a sophisticated instrument.

The derivative question most relevant to artificial intelligence entities is: Who is the perpetrator-via another? There are two possible options: the first is the programmer of the AI software and the second is the user, or the end-user. A programmer of AI software could design the program in such a manner its conduct would be termed as committing offenses via the AI entity. For example: A programmer designs software for an operating robot. The robot is intentionally placed in a factory, and its software is designed to torch the factory at night when no one is there. The robot committed the arson, but the programmer is deemed the perpetrator.

The second person who might be considered the perpetrator-viaanother is the user of the AI entity. While the user doesn't program the software, but he is the ultimate user of the AI entity, including its software, for his own benefit. For example, if the owner of a service robot instructed his dog to attack somebody), then the instructor is held criminally liable[14]

In both scenarios, the actual offense was committed by the AI entity and the programmer or the user did not perform any action conforming to the definition of a specific offense; therefore, they do not meet the actus reus requirement of the specific offense. The perpetration-via-another liability model considers the action committed by the AI entity as if

[13] Ibid
[14]Morrisey v. State, 620 A.2d 207 (Del.1993); Conyers v. State, 367 Md. 571, 790 A.2d 15 (2002)

CC BY-NC 4.0

it had been the programmer's or the user's action. The legal reasoning for construing liability is the instrumental usage of the AI entity who is an innocent agent. The mental attribute required for the imposition of criminal liability is not attributed to the AI entity thereby not meeting the criteria of committing offense[15]. When programmers or users use an AI entity as an instrument, the commission of an offense by the AI entity is attributed to them. The internal element required in the specific offense already exists in their minds.

The perpetration-via-another liability model has many drawbacks that makes it unsuitable when an AI entity decides to commit an offense based on its own accumulated experience or knowledge.[16]Further, this model would fail in a situation where the software of the AI entity was not designed to commit the specific offense, but was committed by the AI entity nonetheless. Also, this model is also not suitable when the specific AI entity functions not as an innocent agent, but as a semi-innocent agent. However, despite the shortcomings the perpetration-via-another liability model might be suitable when a programmer or user makes instrumental usage of an AI entity, but without using the AI entity's advanced capabilities. The legal result of applying this model is that the programmer and the user are criminally liable for the specific offense committed, while the AI entity has no criminal liability whatsoever.[17]

B. The Natural-Probable-Consequence Liability Model: Foreseeable Offenses Committed by Artificial Intelligence Entities

The second model of criminal liability is based on the presumptionof deep involvement of the programmers or users in the AI entity's daily activities, but without any intention of committing any offense via the AI entity.

The first model works on the assumption that the programmers or users have the mens rea, the criminal intent to commit an offense via the instrumental use of some of the AI entity's capabilitiesThe legal situation is different in these cases. In these cases, the programmers or

[15]State v. Fuller, 346 S.C. 477, 552 S.E.2d 282 (2001); Gallimore v. Commonwealth, 246 Va. 441, 436 S.E.2d 421 (1993).
[16] Gerstner M.E.: Comment, Liability Issues with Artificial Intelligence Software, 33 Santa Clara L. Rev. 239. Available at: http://digitalcommons.law.scu.edu/lawreview/vol33/iss1/7 (1993)
[17]Sayre, F. (1930).Criminal Responsibility for the Acts of Another. *Harvard Law Review*, 43(5), p.689.).

users had no knowledge of the committed offense; they had not planned it, and had not intended to commit the offense using the AI entity.

According to the second model, a person can be held accountable for an offense, if that offense is a natural and probable consequence of that person's conduct.[18] Originally, the natural-probable-consequence liability model was used to impose criminal liability upon accomplices, when one committed an offense, which had not been planned by all of them and which was not part of a conspiracy. The established rule prescribed by courts and commentators is that accomplice liability extends to acts of a perpetrator that were a "natural and probable consequence" of a criminal scheme that the accomplice encouraged or aided. Natural-probable-consequence liability has been widely accepted in accomplice liability statutes and codifications.

Natural-probable-consequence liability seems legally suitable for situations in which an AI entity committed an offense, while the programmer or user had no knowledge of it, had not intended it, and had not participated in it, thus not attaching the mensrea element to them[19].The natural-probable-consequence liability model requires the programmer or user to be in a negligent mental state, not more. Programmers or users are not required to know each and every result of their conduct or about any forthcoming commission of an offense as a result of their activity, but are required to know that such an offense is a natural, probable consequence of their actions.

However, the legal consequences of application of the natural-probableconsequence liability model to the programmer or user can differ in two different types of factual cases[20]. The first type of case is when the programmers or users were negligent while programming or using the AI entity but had no criminal intent to commit any offense. The second type of case is when the programmers or users programmed or used the AI entity knowingly and willfully in

[18] Hallevy, G. (2010). The Criminal Liability of Artificial Intelligence Entities. *SSRN Electronic Journal*, pp.179-183.
[19]*21 Santa Clara High Tech. L.J.*. [online] Available at: http://digitalcommons.law.scu.edu/chtlj/vol21/iss1/1 (2004) [Accessed 24 Jan. 2019].
[20]Tuthill G.S(1991) Legal Liabilities and Expert Systems, AI Expert

CC BY-NC 4.0

order to commit one offense via the AI entity, but the AI entity deviated from the plan and committed some other offense, in addition to or instead of the planned offense.[21]

The first type of case is one of pure negligence for which mens rea is not needed to be proven according to this model. The programmers or users negligently acted or failed to act; therefore, there is no reason why they should not be held accountable for an offense of negligence, if there is such an offense in the specific legal system.

The second type of case resembles the basic idea of the naturalprobable-consequence liability in accomplice liability cases. The dangerousness of the very association or conspiracy whose aim is to commit an offense is the legal reason for imposing more severe accountability upon the cohorts.

The question that still needs to addressedis: What is the criminal liability of the AI entity itself when this model is applied? In fact, there are two possible outcomes. If the AI entity acted as an innocent agent, without knowing anything about the criminal prohibition, it is not held criminally accountable for the offense it committed. Under such circumstances, the actions of the AI entity were not different from the actions of the AI entity under the first model (the perpetration-via-another liability model). However, if the AI entity did not act merely as an innocent agent, then, in addition to the criminal liability of the programmer or user pursuant to the natural-probable-consequence liability model, the AI entity itself shall be held criminally liable for the specific offense directly[22]. The direct liability model of AI entities is the third model, as described hereunder.

C. The Direct Liability Model: Artificial Intelligence Entities as being Tantamount to Human Offenders

The third model does not assume any dependence of the AI entity on a specific programmer or user. The third model focuses on the AI entity itself[23]. As discussed above, criminal liability for a specific offense is mainly comprised of the external element (actus reus) and

[21] Kingston J. (2012), Tacit Knowledge: Capture, Sharing, And Unwritten Assumptions. Journal of Knowledge Management Practice, Vol. 13, No. 3.
[22] Warner E.: Expert Systems and the Law. In Boynton and Zmud (eds) Management Information Systems, Scott Foresman/Little Brown Higher Education, Glenview Il, pp. 144-149 (1990)
[23]N.P. Padhy, Artificial Intelligence and Intelligent Systems 3 (Oxford University Press 2005).

CC BY-NC 4.0

the internal element (mens rea) of that offense. Any individual credited with the two components of the explicit offense is considered criminally responsible for that explicit offense[24]. No other criteria are required so as to enforce criminal liability, the existence of the external element and the internal element required to impose liability for the specific offense is quite enough. In order to impose criminal liability on any kind of entity, the existence of these elements in the specific entity must be proven. When it has been proven that a person committed the offense in question with knowledge or intent, that person is held criminally liable for that offense.

When it has been proven that a person committed the offense in question with knowledge or intent, that person is held criminally liable for that offense. The relevant questions regarding the criminal liability of AI entities are: How can these entities fulfill the requirements of criminal liability? Do AI entities differ from humans in this context?

An AI algorithm might have many features and qualifications that exceed those of an average human, but such features or qualifications are not required in order to impose criminal liability. When a human or corporation fulfills the requirements of both the external element and the internal element, criminal liability is imposed. If an AI entity is capable of fulfilling the requirements of both the external element and the internal element, and, in fact, it actually fulfills them, there is nothing to prevent criminal liability from being imposed on that AI entity. Generally, the fulfillment of the external element requirement of an offense is easily attributed to AI entities. As long as an AI entity controls a mechanical or other mechanism to move its moving parts, any act might be considered as performed by the AI entity. Thus, when an AI robot activates its electric or hydraulic arm and moves it, this might be considered an act, if the specific offense involves such an act.

For example, in the specific offense of assault, such an electric or hydraulic movement of an AI robot that hits a person standing nearby is considered as fulfilling the actusreus requirement of the offense of assault[25]. When an offense might be committed due to an omission, it is even simpler. Under this scenario, the AI entity is not required to act at all. Its

[24]William M. Clark & William L. Marshall(1967), LAW OF CRIMES 1-2, 7th ed.
[25] Lawrence B. Solum(1992), Legal Personhood for Artificial Intelligences, 70 N.C. L. REV. 1231.

CC BY-NC 4.0

very inaction is the legal basis for criminal liability, as long as there had been a duty to act. If a duty to act is imposed upon the AI entity, and it fails to act, the actusreus requirement of the specific offense is fulfilled by way of an omission. Attributing the internal element of offenses to AI entities is the real legal challenge in most cases.

Attributing the mental element differs from one AI technology to the other. Most cognitive capabilities developed in modern AI technology are immaterial to the question of the imposition of criminal liability. Creativity is a human feature that some animals possess, but creativity is a not a requirement for imposing criminal liability. Even the least creative persons are held criminally liable. The only mental requirements needed in order to impose criminal liability are knowledge, intent, negligence, etc., as required in the specific offense and under the general theory of criminal law.

CONCLUSION

So it appears that the question of whether AI system can be held legally liable relies upon three components:

• The restrictions of AI frameworks, and whether these are known and imparted to the buyer; Since AI systems have both general and specific limitations, legal cases on such issues may well be based on the specific wording of any warnings about such limitations.

• Whether an AI framework is a product or service; There is an ambiguity with respect to classification of AI in law; different commentators offer different views.

• Whether the offense requires a mensrea or is a strict liability offense; It seems unlikely that AI programs will contravene laws that require knowledge that a criminal act was being committed; but it is very possible they might contravene laws for which ‘a reasonable man would have known’ that a course of action could lead to an offence, and it is almost certain that they could contravene strict liability offences

If an AI system is held liable, the question arises of whether it should be held liable as an innocent agent, an accomplice, or a perpetrator depending on which of Hallevy’s three models is applied.

CC BY-NC 4.0