INTERNATIONAL JOURNAL OF RESEARCH AND ANALYSIS VOLUME 5 ISSUE 2 ISSN 2347-3185 2018

FROM SCI-FI TO STEEL COLLAR CRIMES

**SONAL SODHANI & HIMANSHU CHITLANGIA

1. Introduction

Increasing importance and utilization of Artificial Intelligence have been perceived and recognized all through world. From all kind of electronic equipment to robots, it is playing a significant part. Its optimum utilization, might be seen in computer. The innovative and techno-friendly world is graving rapidly the area of typical human activities either in form of artificial intelligence loaded machine or robots. Since then Artificial Intelligence (hereafter called as AI) entities have turned into an indispensable part of present day human life, working considerably more adroitly and sophisticatedly than other daily tools. But, could they become dangerous and harmful? Actually, they already have become so.

AI an asset or a weapon? And if it turns around to be a weapon then who shall be liable for the same. The weapon itself or the user of the weapon?

In 1950, Issac Asimov[1] laid down three fundamental laws of robotics in his sci-fi magnum opus:

i) A robot shall not harm or injure a person or, through inaction, permit an individual to be harmed;

ii) A robot must comply with the commands or orders given to it by a human, except where such requests would be in conflict with the first law;

iii) A robot may protect its own particular existence, as long as such insurance or protection does not conflict with the first and second laws.[2]

The laws of Asimov of robotics is only in relation to robots. AI programming software installed in a robot would not be subject to Asimov's laws, even if these laws had any lawful and legal importance. The fundamental question in this setting is which kind of laws or morals are right and who is to decide such cases. Criminal law encapsulates the most powerful legal social control in present day civilisation. But the

*Student, New Law College, Bharti Vidhyapeeth Deemed University, Pune.
**PGDM, Xavier Institute of Social Service, Ranchi.

[1] Issac Asimov, *I. Robot* (1950).

[2] Rae Johnston, *Robots could be given "Human Rights" as early as next month,* GIZMODO, https://www.gizmodo.com.au/2017/01/robots-could-be-given-human-rights-as-early-as-next-month/ (on Mar. 15, 2018, 11:38am).

fact is, AI entities are not considered to be subject to the law, especially to criminal law.[3] The primary question with respect to AI entities is – does the developing intelligence of AI entities subject them to legal, social control, as some other legal entities, in a similar manner as corporations? The author tries to work out a legal solution to the problem of criminal liability of AI entities.

2. What is an Artificial Intelligence entity?

For a few years, there has been significant discussion and debate about the very essence of an AI entity.[4] Futurologists have proclaimed the emergence of another new species, *machina sapiens*, which will share the human place as savvy intelligent creatures on earth. Critics have contended that a "thinking machine" is an oxymoron. Machines, including computers, with their foundations of cold rationale, can never be as insightful or inventive as humans.

Roger C. Schank, an American artificial intelligence scholar and theorist, in one of his articles, set down five attributes or characteristics that one would anticipate that a wise, intelligent entity will have. These attributes are enlisted below:

i) **Communication**

You can speak with your pet puppy, however you can't talk about material science or physics with it. You can't speak with a stone by any means. You can't discuss quantum mechanics with a 3 year old. In other words, the easier it is to communicate with an entity, the more intelligent it appears.

ii) **Internal Knowledge**

An intelligent being is expected to know his identity, who he is and what he thinks.

iii) **External Knowledge**

An intelligent entity is expected to know as well as think about the outside world, to find out and learn about it, and use that information. Consequently, intelligent entities must have a capacity to see new encounters and experiences in terms of old ones.

iv) **Intentionality**

Goal-driven conduct implies knowing when one needs something and knowing a method or a plan to get what one needs. More often than not, a presumed correspondence exists between the intricacy of the objectives that an entity has and the sheer number of plans which an entity has access to achieve these objectives.

[3] Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*. 70 N.C. L. REV. 1231 (1992).
[4] Terry Winograd, *Thinking Machine: Can There Be? Are We?* , THE FOUNDATIONS OF ARTIFICIAL INTELLIGENCE, 167 (Derek Partridge and Yorick Wilks eds., 2006).

Thus, a tree has none or next to none of these plans and objectives, a puppy has fairly more, and a man has many; very intelligent people presumably have more.

v) **Creativity**

An intelligent entity is relied upon to have some degree of creativity. In this specific context, creativity implies the capacity or ability to take alternative action when the initial action comes up short and fails. A fly tries to leave a room and bumps into a windowpane, it keeps on repeating the same pointless and futile conduct. But, at the point when an AI robot bumps into a window, it tries to leave using the door.[5]

Some 21th century type of AI entity has significantly more attributes that empower them to act in form of more refined and sophisticated ways.[6]

A robot can be designed to mimic and imitate the physical capabilities of an individual person and these abilities can be improved and enhanced. A robot is capable of being physically faster and stronger than a human being. The AI programming software installed in it additionally empowers the robot to calculate numerous complicated computations and calculations faster and simultaneously, or to "think" faster. An AI entity is capable of learning and gaining experiences, and experience is a useful method for learning. Every one of these traits and attributes creates the essence of an AI entity. AI robots and AI programming software are utilized as a part of an extensive range of applications in industry, military administrations, medical services, science and even games and amusements.

3. Criminal liabilities of Artificial Intelligence entities

In 1981, a 37-year-old Japanese worker of a motorcycle processing factory was killed by an artificial intelligence robot working close to him. The robot mistakenly identified the worker as a risk to its main mission, and calculated that the most proficient approach to dispose of and eliminate this danger was by pushing him into an adjacent operating machine. By using its very powerful arm, the robot smashed the astounded employee into the operating machine, killing him right away, and after that continued its duties with nobody to meddle or interfere with its main mission. [7]

Is this murder? Is the robot a killer?

---

[5] Roger C. Schank, *What Is AI, Anyway?* , AI MAG., Winter 1987, at 59.

[6] In November, 2009, during the supercomputing conference in Portland Oregon, IBM Scientists and other announced that they succeeded in creating new algorithm named “Blue Matter”, which possesses the thinking capabilities of a cat. Chris Capps, *“Thinking” Supercomputer Now Conscious as a Cat*, http://www.unexplainable.net/artman/publish/article_14423.shtml (on Mar. 28, 2018, 5:27pm).

[7] Yueh-Hsuan Weng, Chien-Hsun Chen & Chuen-Tsai Sun, *Toward the Human-Robot Co-Existence Society: On Safety Intelligence for Next Generation Robots*, 1 INT. J. SOC. ROBOT 267, 273 (2009).

Subsequent to having effectively established what systems of AI are, comes a question under consideration – can artificially intelligent machines be held liable for violations conferred and crimes committed by them?

To decide the criminal liability of AI entities for a particular offence carried out for specific point of time and space is a challenging task in legal field. Two components are fundamental for imposing criminal liability on any individual – first, external or factual component i.e. criminal conduct (actus reus), while the other component is internal or mental component, i.e. knowledge or general intention vis-à-vis the conduct (mens rea). On the off chance that any of them is absent, no criminal liability can be imposed.

The actus reus requirement is communicated primarily by the demonstrations of an act or its omission.[8] The mens rea requirement can be communicated and expressed through various methods, such as, by knowledge, by specific intent, by negligence or even by strict liability.

No other criteria or capabilities are required to impose criminal liability, neither from human beings, nor from some other kind of entity, including corporations and AI entities.[9] An entity may have further abilities, be that as it may, in order to impose criminal liability; the presence of actus reus and mens rea in the particular offence is quite enough. A honeybee is capable of acting, but the aspect of wrong intention is missing, hence, it is incapable for defining the mens rea requisite; subsequently, a bee sting bears no criminal obligation or liability. A parrot is capable of repeating words it hears, but it has no mind to analyse the nature of what it is saying, hence, it is incapable for defining the mens rea requisite for libel.

In order to impose criminal liability on any kind of entity, it must be demonstrated and proven that the above two components existed. When it has been demonstrated that a man carried out the criminal act intentionally having knowledge of the same or with criminal intent, that individual is held criminally liable for that offence. The significant question concerning the criminal liability of AI entities is – How would Artificial Intelligence entities be able to satisfy the two necessities of criminal liabilities?

Gabriel Hallevy, an associate professor in the Ono Academic College of Israel, in one of his well-read papers[10], introduced three possible models for imposition of criminal liabilities on AI entities, and the same shall be discussed below:

[8] Walter Harrison Hitchler, *The Physical Element of Crime*, 39 DICK. L. REV. 95 (1934).
[9] Joshua Dressler, CASES AND MATERIALS ON CRIMINAL LAW, 980-81, (2007).
[10] Gabriel Hallevy, *Virtual Criminal Responsibility*, SSRN, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=183562 (on Apr. 1, 2018, 10:57am).

- The Perpetration-via-another liability model;
- The Natural-Probable-Consequence liability model; and
- The Direct liability model.

4. The Perpetration-via-another Liability Model

This first model does not consider the AI entity as having any human characteristics. The AI entity is viewed as an innocent agent.[11] Accordingly, due to that legal perspective, a machine is a machine, and is never human. In any case, one can't disregard an AI entity's capabilities, as specified previously. An AI entity has a brain that functions, a mind that learns.

In accordance with this model, these capabilities are insufficient to regard the AI entity a culprit of an offence. These abilities resembles the parallel capacities of a mentally limited individual, for example, a kid, a man who is mentally unsound, or one who does not have a criminal state of mind.

Lawfully, when an offence is carried out by an innocent agent (a child, a man who with unsound mind, or one who does not have a criminal state of mind) that individual is criminally liable as a perpetrator-via-another.

In such cases, the intermediary is viewed as a mere instrument, while the party orchestrating the offence (perpetrator-via-another) is the real culprit as a principal in the first degree and is considered responsible for the conduct of the innocent agent.

The subordinate question with respect to artificial intelligence entity is – Who is the perpetrator-via-another? There are two candidates – the first is the developer of the AI software and the second is the user, or the end-user. A developer of AI software may outline a program to commit offence by means of the AI entity. For instance, a developer designs a software for a working robot. The robot is deliberately placed in a factory, and its software is designed to torch the manufacturing plant during the evening when nobody is there. The robot committed arson, but the developer is regarded the culprit. The second individual who may be viewed as the perpetrator-via-another is the user of the AI entity. The user did not program the product, but rather he utilizes the AI entity, including its software, for his own personal advantage. For instance, a client buys a servant-robot, which is designed to execute any request given by its master. The robot recognizes the particular user as the master, and the master orders the robot to assault any intruder of the house. The robot executes the request precisely as requested. This is similar to the

[11] Lawrence B Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231 (1992).

situation when a person orders his dog to attack any trespasser. The robot attacked, but the user is regarded the culprit. In both the situations, the real offence was conferred by the AI entity. The programmer or the user did not perform any activity fitting in with the meaning of a particular offence; in this way, they don't meet the actus reus requisite of the particular offence. The perpetrator-via-another liability model considers the activity conferred by the AI entity as though it had been the developer's or the user's action. When developers or users use an AI entity instrumentally, the commission of an offence by the AI entity is credited to them.

This model does not attribute any mental capacity, or any human mental ability, to the AI entity. As per this model, there is no lawful difference between an AI entity and a knife or an animal.

Limitations

This model is not suitable for cases where-

- An AI entity chooses to confer an offence based on its own accumulated understanding or knowledge.
- The software of the AI entity was not designed to confer the specific offence, but was perpetrated by the AI entity nonetheless.
- The specific AI entity functions as a semi-innocent agent and not as an innocent agent.

The lawful result of applying this model is that the developer and the user are criminally liable for the particular offence perpetrated, while the AI entity has no criminal obligation at all.

5. The Natural-Probable-Consequence Liability Model

The second model of criminal liability assumes profound contribution of the developers or users in the AI entity's every day activities, yet with no intention of committing any offence by means of the AI entity.

One scenario: amid the execution of its every day tasks, an AI entity confers an offence. The developers or users had no knowledge of the offence until the point that it had just been committed; they didn't plan to commit any such offence, and they didn't even participate in any part of the commission of that particular offence. One example of such a situation: an AI robot or software, which is designed to work as a programmed automatic pilot which has been customized to protect the mission as a part of the mission of flying the plane. Amid the flight, the human pilot activates the automatic pilot (which

is the AI entity), and the program is initiated. At some point, after initiation of the programmed pilot, the human pilot sees an approaching storm and tries to abort the mission and come back to base. The AI entity deems the human pilot's activity as a threat to the mission and takes action to eliminate that danger. It may remove the air supply to the pilot or activate the ejection seat, and likewise. Subsequently, the human pilot is killed by the AI entity's activities.

Clearly, the developer had not intended to kill anybody, but the human pilot was killed because of the AI entity's activities, and these actions were done according to the program.

The first model assumes mens rea, the criminal intent of the developers or users to perpetrate an offence by means of the instrumental use of a portion of the AI entity's abilities. This isn't the case here. In these cases, the developers or users had no knowledge of the conferred offence; they had no intention to commit the offence using AI entity. For such cases, the second model may make a reasonable lawful response. This model depends on the ability of the developers or users to anticipate the potential commission of offences.

As indicated by the second model, a man may be considered responsible for an offence, if that offence is a natural and probable consequence of that individual's action.

The developers and users of an AI entity, who ought to have known about the probability of the imminent commission of an offence, are criminally liable for the particular offence, despite the fact that they didn't really know about it. This is the fundamental legal basis for criminal obligation in cases of negligence. Negligence is an oversight and omission of awareness or knowledge.

The question still remain – What is the criminal obligation or liability of AI entity itself when the Natural – Probable – Consequence model is connected? There might be two conceivable outcome –

a) When AI entity acted as an innocent agent, with not much knowledge, it isn't responsible for conferred offence.
b) When AI entity did not act as an innocent agent, then along with criminal liability of developer or user (as Natural-Probable-Consequence Model), the AI entity itself should also be held criminally liable for the particular offence directly.

6. The Direct Liability Model

Recently, Facebook had to abandon an experiment after two artificially intelligent programs appeared to be chatting to each other in a strange language which only they

comprehended and understood.[12] The two chatbots made their own specific changes to English that made it pretty easier for them to work – however, which remained mysterious to the humans who are supposed to look after them. The robots were instructed to work out how to negotiate amongst themselves, and enhance their trading as they went along. However, they were not told to use comprehensible English, enabling them to make their own "shorthand", according to the researchers.[13]

In such cases, if some mishap happens then who is to be held liable, as the act done was neither commanded by the user nor did the developer made the program to do such an act. Also to be noted, AI has the ability to write and rewrite its own codes.

The third model does not assume any reliance of the AI entity on any developer or client. The third model focuses on the AI entity itself.

Criminal obligation for a particular offence predominantly includes actus rea and mens rea. Nothing other than this is required to impose criminal liability on a person. To impose criminal obligation on any entity, the presence of these components in the particular entity must be proven. When it has been proved that a man perpetrated the offence being referred to with intention, that individual is held criminally liable for that offence.

The important query in regards to the criminal liability of AI entities is – How can these entities satisfy the necessities of criminal liability? On the off chance that an AI entity is able to satisfy the prerequisites of a crime i.e. actus rea and mens rea, and there is nothing to keep criminal liability from being imposed on that AI entity.

Generally, the satisfaction of the actus rea of an offence is easily ascribed to AI entities. Ascribing the mens rea of offences to AI entity is the real lawful challenge. Attributing the mental component differs from one AI technology to the other. The main mental necessity required to impose criminal liability are knowledge, intent, negligence, and so on, as required in the particular offence and under the general hypothesis of criminal law.

Knowledge is characterized as sensory reception of factual data and the comprehension and understanding of that data. Most AI frameworks are equipped for such reception. Sensory receptors of sights, voices, physical contact, touch, and so forth, are basic in most AI systems. These receptors transfer the data to central processing units that analyse the data.

[12] Andrew Griffin, *Facebook Artificial Intelligence Robots Shut Down After They Start Talking To Each Other In Their Own Language*, (July 31, 2017, 16:10 BST), https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html (on Mar. 20, 2018, 11:21pm).
[13] Ibid.

Specific intent is the most important of the mental element prerequisites. An AI entity may be programed to have a reason or a purpose and to take actions to accomplish that aim. This is specific intent.

Why should an AI entity that satisfies all components of an offence be absolved from criminal obligation?

Most AI algorithms are capable to analyse – good and bad. Likewise, when an AI algorithm works appropriately, there is no reason behind it not to use every one of its abilities to analyse the data received through its receptors.

When all components of particular offences have been proved in an AI entity, there is no reason to avoid imposing criminal obligation upon it for that offence. This criminal liability of AI entity does not supplant the criminal liability of the developers or users.

Not only the positive elements that constitute to the crime, but also the elements that is considered to be general defence in criminal law, might be attributed to AI entities. A defence of insanity may be brought up in connection to a failing of AI algorithm, when its scientific capacities gets corrupted because of that malfunctioning. Similarly, the impact of electronic viruses which affects the working system of AI entity may be viewed as parallel to the impact of intoxicating substances on people.

It may be summed up that the criminal liability of an AI entity as indicated by direct liability model is similar to the significant criminal liability of a human. Though, a few changes are necessary, however substantively, it is same.

7. Punishment aspects

The end of this reaction is punishment. Let us assume an AI entity is criminally liable, it is prosecuted, tried, and sentenced. After the conviction, the court sentenced that AI entity, for instance, a year of imprisonment. How can an AI entity practically serve such a sentence? How can arrest be made? By what means would death sentence or fine be influenced against AI entities?

Similar legal issues were raised when the criminal obligation of corporations was perceived and recognised.

The answer was simple and legally applicable. When a punishment can be imposed on a corporation as it is on humans, it is imposed without change. When the court adjudicates a fine, the corporation pays the fine similar to the way a human pays the fine. Nonetheless, when punishment of a corporation can't be done in the same manner as done with humans, an alteration and adjustment is required. Such is the legal situation and circumstance vis-à-vis AI entities.

Certain adjustments that can be made regarding AI entities –

- **Death Penalty –** The life of the AI entity is an independent existence and presence as an entity. Considering capital punishment, this might be accomplished by deletion of software from AI entity. The deletion annihilates the autonomous presence of the AI entity and is tantamount to capital punishment.
- **Imprisonment –** Considering the nature of a sentence of imprisonment, the practical action that may accomplish a similar impact as imprisonment is when AI entity is to put out of use for a decided period. Amid that period, no activity identifying with the AI entity's freedom is permitted, and consequently its liberty and freedom is confined.
- **Community Service –** The importance of the community service for human is mandatory commitment of work towards the community. An AI entity can be engaged in as labourer in numerous area. When an AI entity works in a factory, its work is done so as to ease and facilitate the workers task. Similarly, AI substance works for benefit of private individual and community too. The same might be imposed by sentence.
- **Fine –** AI entities have no money or property of their own or bank account. Fine imposed on an AI entity may be collected in the form of work and labour for the benefit and advantage of State or community.

8. Conclusion

It's a war between an inventor and its invention. The basic idea behind this is empowering the invention with a power which an inventor can control.

We humans have successfully given birth to a complete new species. We have made ourselves god but to be true this god now is possessing an extinction threat by its own creation. Have we taken it too far? The nature's law have been changed and so will the rules of extinction. An environment of co-existence is still working on the principles of probability. The rapid development in AI technology requires current legal solutions that protect society from conceivable harms and damages inherent in technologies. It is the key function of a legal system to preserve social order for the benefit and welfare of the society. It is clear that the oblivion in law need to be settled with regards to a system that is running further and further ahead of human control, without many people having a clear idea on who is in-charge, or whose responsibility it might be when something goes wrong.