Legal Personality of Artificial Intelligence: Necessity and the Imposition of Liability

Tanisha Sudarshan

Abstract:

Advancement of technology has been a constant wheel of motion ever since the earliest times. The Industrial Revolution is considered as one of the biggest achievements of mankind that follows the invention of fire. Presently, in the Twenty-first Century, Advancements in Computer Technology has grown leaps and bounds and society is faced with a new challenge of accepting Artificial Intelligence (AI) Technology. While the Technical challenges in incorporating this technology seamlessly into human lives are being dealt by engineers, there stands a bigger concern for lawyers and policy-makers with regards to the jurisprudential status of such beings. The global scenario in this regard has been divided into two camps, one which is apprehensive of granting legal status to AI, while other has gone ahead and recognized AI as legal persons, popular instances being that of Saudi Arabia granting citizenship to the Humanoid Robot Sophia in October 2017. Japan has quickly followed by granting official residence to a chatbot designed to be a seven-year old boy, Shibuya Mirai. Such instances give rise to questions about the necessity of a personality status to AI. The most distinctive debate is that whether AI are to be treated as machines or should they be accorded a separate status. By answering this question, it is possible to attribute liability arising out of the acts of such AI-based computers and machines. This paper aims to deal with the idea of granting of personhood to AI and thereby, determining the liability in such instances.

Keywords: Artificial Intelligence, Jurisprudential status, Liability, Personality, Robot.

INTRODUCTION

Computer technology has become an inevitable extension of human lives and has strongly rooted itself. Robots and self thinking machines have been portrayed in popular culture and cinemas for a long time now. In recent times, the concept of Artificial Intelligence has gained popularity and has grown exponentially. Prominent nations across the world have made declarations recognising humanoid robots and chatbots as citizens and residents. While the technology has developed in a steady pace, the law remains uncertain to integrate such instances in the functioning of the social and political spheres. This paper is an attempt to understand the nature of Artificial Intelligence (AI), the legal standing of AI and the means to resolve any legal incidents. The second section of the paper deals with the technical know-how in relation to AI. The third section attempts to philosophise the nature of AI in relation to the human intelligence. The fourth

section emphasises on the importance of liability with relation to a right, thus highlighting the importance of realising liabilities to a wrong. The fifth section is the researcher's attempt at devising various means of accruing liability and can be further divided into three distinct sections. The conclusion summarises the paper and also suggests different tools to deal with AI on a large scale.

INTRODUCTION TO ARTIFICIAL INTELLIGENCE

The term Artificial Intelligence (AI) is referred to those form of computer programs that perform tasks that require human cognition or thinking. Luger and Stubblefield define artificial intelligence as the branch of computer science that is concerned with the automation of intelligent behaviour.[1] This highlights the question as to how a computer is said to possess intelligence. To determine whether a computation falls under the category of Artificial Intelligence or not, the Turing Test given by Alan Turing is used.

The Turing Test was proposed to test for human level intelligence using an "imitation game". The original version is played by three people, a man (A), a woman (B) and an interrogator (C), who may be of either sex. The interrogator stays put in a room apart from A and B. The object of this test is for C to determine which of the two participants is a man and the other a woman. The interrogator may use a teletype to question the two participants and conclude based on the answers thus given. It is A's object to mislead the interrogator while B has to aid the interrogator. To test for determination of AI, A is replaced by the computer. The computer is said to possess AI if it is successful in fooling the interrogator.[2]

In popular science-fiction, robots that possess AI are portrayed to possess almost human like behaviour, which is a characteristic of advanced AI systems. But AI systems are of different types, predominantly distinguishable into four broad categories:

- Reactive Machines: This type of AI system is rudimentary and basic. It is only reactive and does not possess the ability to retain experiences as memories or retrieve and use past experiences. It involves the computer reacting to the stimulus. The AI can be said to lack individuality. IBM's Deep Blue is an example for this type. The chess playing super computer can identify the pieces on a chess board and know how each moves. It can make predictions about what moves might be next for it and its opponent. And it can choose the most optimal moves from among the possibilities.
- Limited memory Machines: These machines have the ability to look into the past and react accordingly. Self Driving Cars are a good example under this type. They observe other cars' speed and direction, which is cannot be done momentarily, but rather, requires identifying specific objects

[1] N. P. Padhy, Artificial Intelligence and Intelligent Systems, at 7 (1st ed. 2005).
[2] Larry Hauser, Artificial Intelligence, Internet Encyclopaedia of Philosophy (Aug 26, 2018, 09:20 a.m.) https://www.iep.utm.edu/art-inte/#H2

and monitoring them over time. But these simple pieces of information about the past are only transient. They aren't saved as part of the car's library of experience it can learn from, the way human drivers compile experience over years behind the wheel.

- Machines that think: This class of computers not only forms representations about the world, but also about other agents or entities in the world. The fundamental goal is to understand that people, creatures and objects in the world can have thoughts and emotions that affect their own behavior.
- Self aware machines: The final step of AI development is to build systems that can form representations about themselves. Such machines are said to possess consciousness. Conscious beings are aware of them, know about their internal states, and are able to predict feelings of others. This kind very closely resembles human intelligence and behavior.[3]

PERSONHOOD TO AI AND THE PHILOSOPHICAL UNDERPINNINGS

In October 2017, Saudi Arabia became the first country to give citizenship to a humanoid robot, Sophia, created by David Hanson from Hanson Robotics. Sophia is in the infancy of Artificial Intelligence. At the Future Investment Initiative held at Riyadh, the Kingdom announced that they would be granting citizenship to Sophia. There has been no elaboration on the rights and duties that Sophia possesses as a citizen.[4] In a similar bid, Japan has made a chatbot named Shibuya Mirai as a resident in Tokyo.[5] Such declarations by prominent nations put a burning question about the status of such AI bots.

The Black's Law Dictionary elucidates the term "citizen" as "a member of a free city or jural society, possessing all the rights and privileges which can be enjoyed by any person under its constitution and government, and subject to the corresponding duties." [6] Thus, recognition as a citizen or a resident implies attribution of personality to such robots and chatbots because a legal right is vested in a person, who may also be distinguished as the owner of such a right. Further, "a person is any being whom the law regards as capable of rights and duties. Any being that is so capable is a person. Persons are the substance of which rights and duties are the attributes."[7] It is of course noted that non-humans are also accorded persons, but granting of personality to AI is a crucial discussion owing to their nature. While politically, personality has

[3] Arend Hintze, Understanding the four types of AI, from reactive robots to self-aware beings, The Conversation (Aug 27, 2018, 00:45 a.m.) https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

[4] Zara Stone, Everything you need to know about Sophia, the World's First Robot Citizen, Forbes (Aug 27, 2018, 13:49 p.m.) https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#439f9c4346fa

[5] Agence France- Presse, Tokyo's 'AI' boy first bot to gain residency, The Hindu, November 04, 2017 (Aug 27, 20118, 14:02 p.m.) https://www.thehindu.com/news/international/tokyos-ai-boy-first-bot-to-gain-a-residency/article19982384.ece

[6] Black's Law Dictionary Free Online Legal Dictionary (2nd ed.) (last accessed on 10/09/2018, 12:51 p.m.) https://thelawdictionary.org/citizen/

[7] P J Fitzgerarld, Salmond on Jurisprudence, at 62 (South Asian ed. 2016).

been attributed to AI, it is prudent to examine the nature of AI before legally acknowledging such declarations.

Intelligent behavior in computers is possible with algorithms, a specific set of operations and procedures and decisions which guarantee to yield correct results.[8] Such algorithms form the basis of machine learning, whereby instead of detailed programming, computers are given general goals, which then, like humans, improve with experience. [9]This body of instructions and goal setting, according to American Philosopher Hubert Dreyfus, can be understood as the Background. He follows German Philosopher Martin Heidegger and argues that skillful coping is the fundamental phenomenon and that the Background is presupposed by it. This background is not an explicit body of facts and rules because; a programmer cannot code all possible facts which are relevant in a multiplicity of situations. If the programmer were to do so, it would then lead to an infinite regress of facts about other facts. In order to avoid this regress, it is but required that the background is not completely and explicitly represented.[10] The machine uses this set of background facts and learns to improve the output by mimicking natural selection, through evolutionary algorithms.[11]

However, human intelligence function is not mainly based on putative background facts, but also based on brain processes, with distinct cognitive abilities. It is here that human intelligence can be differentiated from that of AI. Human intelligence can also be understood as the result of consciousness, an idea that is abstract from machines. This relates to the mind-brain problem, that is, construction of conscious intelligence by physical systems like the brain. The dichotomous concepts of dualism and materialism act as tools to solve this problem. The 17th century French Philosopher, Rene Descartes expounds the theory of Dualism, where the mind and brain are two distinct and separate phenomena, and conscious experience is non-physical and beyond the scope of physical sciences. Materialism on the other hand asserts that mind and body are both physical media and that by understanding the body and brain, the understanding of the mind will follow.[12] These philosophical ideas tend to overlook one another, thus failing to clearly define consciousness.

However, cognitive psychologist Steven Pinker has suggested that consciousness can be broken down into Self knowledge, access to information and sentience. According to Pinker, an intelligent being has the most accurate information of itself, making self-knowledge no more mysterious than perception or memory. The ability to report on the content of information without the capacity to report on the build-up of such content forms the basis of access awareness. This accessing of information is limited only to the conscious level and does not surpass to the unconscious. Lastly, sentience, also known as *qualia*, refers to subjective experience,

[8] N. P. Padhy, supra, at 10
[9] Pedro Domingos, AI Will Serve Our Species, Not Control It: Our Digital Doubles, 319 Sci. Am. 88-93. (20118) EBSCO*host*, search.ebscohost.com/login.aspx?direct=true&db=pbh&AN=131111192&site=ehost-live.
[10] Beth Preston, Heidegger and Artificial Intelligence, 53 PPR 43, 45 (1993)
[11] Supra note 9
[12] Michael S. Gazzaniga et al., Cognitive Neuroscience The Biology of the Mind at 609 (4th ed. 2014)

raw feelings and phenomenal awareness.[13] Since, cognitive psychology uses human anatomy to explain the concept of consciousness; the brain is considered as the seat of intelligence, then in computers, the synonymous position would be attributed to the Central Processing Unit. If Pinker's theory of consciousness is to be applied to computers, it is still possible that the initial two prongs, self-knowledge and access awareness is found in computers. The difference established before human and artificial intelligence comes down to sentience. While this phenomenon cannot be accurately answered, it is evidently a part of human consciousness, which can be connected to Dreyfus's concept of internal representation. In the end, it can be rightly said that any theory of intelligent behavior essentially depends on the internal representation of the background.

Intelligent behavior is also characterized by free will. Certain theorists assert that AI is just the ability to solve problems- as task that does not require free will.[14] A counter-claim to this is Kant's Theory of Free Will and Pure Practical Reasoning. For Kant, reason governs Will and the capacity for reason is bound by the capacity for freedom.[15] He states that when behavior is biologically determined, it is not truly free. Autonomous behavior is not bound by dictates of nature or social convention. To act freely is not to choose the best means to a given end; but to choose the end itself, for its own sake.[16] When Kantian theory is applied to AI, it suggests that the AI is autonomous, since its behavior is not governed by biological dictates, unlike human beings. As stated above, the AI is driven by goals, which is the end in itself.

If AI is backed by Kant's Pure Practical Reason, it then becomes more human than humans, the epitome of Kantian philosophy of will and reason. This is because, AI, unlike humans is not sentient, hence do not deviate from the Categorical Imperative as suggested by Kant. This conclusive reasoning is backed by taking into consideration an instance from the 2004 Hollywood movie, 'I, Robot', the protagonist is rescued by a humanoid robot life-guard, staking the life of a 12-year old girl, who had very little chance of living. The robot here did not get emotionally influenced but calculated the probability of survival of each of them and thus picked the one with high chances, thus achieving its purpose of creation. If the same were to be demonstrated by a human, then he/she might have attempted to save the child, driven by emotions, thus staking the life of the person who had more odds of survival. But merely achieving the ability to reason does not mean that AI should be accorded personhood. While the robot's act was in perfect consonance with Kantian thinking, it is still a far cry from accepted behavior as emotions and sentience is what distinguishes humans from the rest of the species. It can thus be concluded that machines, however intelligent cannot be accorded the same status as humans.

[13] *Id*
[14] Supra note 9
[15] Michel J Sandel, Justice: What's the Right Thing to do? at 108 (reprint ed. 2011)
[16] *Id* 109

THE NECESSITY TO IMPOSE LIABILITY

Citing Saudi Arabia and Japan as an example, it is possible that in future these "Artificial persons" will become a part of the day to day activities in the social sphere, interacting with other persons. This brings up the question on the consequences of their interaction and the legal stand thereof. To understand the relation between citizenship, personality and the rights and duties that spring from such attribution, the analysis by the American Jurist Wesley Newomb Hohfeld is of importance. The vesting of the status of person is to confer certain rights upon the subject. To ascribe a right to one person is to imply that some other person is under a corresponding duty.[17] To say that A has a legal right against B is to say that A is legally protected from B or A has a legal sanction towards B to expect an act in favour of him. So in the circumstance of Sophia or Shibuya Mirai, they are vested with certain rights and that other persons have a duty towards them. The vice versa is also true. It would thus mean that Sophia or Mirai are duty bound towards other persons who have a right against them. Every right or duty involves a *viniculum juris* or bond of legal obligation.[18] This means that without either right or duty, the other ceases to exist.

The non-fulfillment of a duty violates the rights and leads to a circumstance of a wrong. When wrongs are committed, the wrong-doer is bound to rectify this. Liability is the bond of necessity that exists between the wrongdoer and the remedy of the wrong.[19] The concept of liability is the converse of the Latin maxim *ubi jus ibi remedium*, translated as where there is no right, there is no remedy. It is then right to state that when a person commits a wrongful act then he is bound to provide a remedy.

When a wrongful act is committed, the wrongdoer compromises the status quo and he has the legal sanction to restore the status if the injured. The wrong may either be a violation of a *right in personum*, attracting civil liability or a violation of a *right in rem*, attracting criminal liability. The legal sanctions result in remedying the wronged person by either paying tortious damages or by serving penal consequences. Although, AI has a higher accuracy than humans, it is possible that malfunctions occur and this might hamper with another person's rights. Thus, it is preemptive to impose liability on wrongs committed by Artificial Intelligence too.

ACCRUING LIABILITY

The above section has conclusively stated that there is no escape from accruing liability to AI. Since AIs and machines are the creations of humans, it is proposed that there be no grant of an independent status to such machines and the liability that arises thereof, be attached to that of a natural person, the creator or in

[17] P. J. Fitzgerald, supra, at 217
[18] *Id* at 220
[19] *Id* at 349

certain cases, the owner. What follows is an attempt at imposing liability through different modes and drawing parallels to the existing system.

- **Minority and AI**

It is proposed that AI be given a juristic personality. In India, specifically, idols are considered to be juristic persons. An idol is a deity used as an object of worship, which according to precedents, has been given a legal or juristic personality. But, an idol is treated as a minor, on behalf of whom the guardian, usually known as the *shebait* can go on about its business, with respect to holding property.[20] In relation this, the AI can be treated as a minor by the law and the creator or the owner is accrued parent liability.

There are varied principles of law that deals with parental liability. The old principle of law was that guardians were not personally held liable for torts committed by minors under their charge.[21] But the stance that law has with regard to parental liability has changed. In the United States, most states' laws impose civil liability for acts done "willfully or knowingly" and it does not include accidents. The idea behind imposing parent liability is to ensure that parents are more attentive towards their wards. Such imposition acts as a deterrent against poor parenting.[22] Taking minority status of AI into consideration, "willful activities" is questionable, but the principle behind attributing liability can be adopted. Since imposition of liability acts as a deterrent, it will also aid in ensuring that the AI system is not mal-functioning, due to which civil liabilities may arise.

In case of a criminal liability, most cases of parental liability is imposed due to the negligence of the parents. In *Kennedy v. Baird*, the plaintiffs went to the defendant's house where the defendant's son began shooting at them and injured them. The plaintiffs sued the defendant for negligently entrusting his son with a gun. The court of appeals held that it was possible to bring an action for negligent entrustment of a firearm; however, the court affirmed the dismissal of the plaintiffs' claims where there was no evidence that the defendant knew or should have known that his son was likely to commit such a crime.[23]In California, it is considered as a misdemeanor if a parent fails to fulfill his or her "duty to exercise reasonable care, supervision and control over their minor child."[24] The California Supreme Court rejected the contention by a group of taxpayers that such a law is unconstitutional. The court found that the law promoted the state's legitimate interest in addressing adolescent delinquency and gang activity.[25] When minority status is accorded, it incentivizes the owner or creator of such technology to use it for peaceful purposes. It also helps to avoid negligence on part of the creator or the owner.

[20] Yogendra Nath Naskar v. Commissioner of Income Tax (1969) 3 SCR 742
[21] Luchmun Das v Narayan (1871) 3 NWP 191
[22] Janet Portman, Parents' Responsibility for their Child's Actions, Lawyers.com (last accessed on 14/09/2018 at 22:45 p.m.) https://www.lawyers.com/legal-info/criminal/juvenile-law/parents-responsibility-for-their-childs-actions.html
[23] Kennedy v Baird 682 S.W. 2d 377 (1984)
[24] Cal. Penal Code § 272 (1872)
[25] Williams v Garcetti 853 P.2d 507 (Cal. 1993)

- **AI as an agent**

In modern times, it is highly possible that computers are involved in contracting. The prime example of this would be Google's Duplex. The video demo released by Google in early 2018 has an AI application making phone calls to different businesses and make appointments. This is a typical example of a principal-agent relationship, only here, the AI is the agent. The concept of agency involves the agent acting on the instruction of the principal, whose relationship is also governed by a contract. The challenge at hand is to determine whether the nature of relationship between the two is in fact capable of being governed under Contract Law.

To constitute a valid contract, there must be two or more separate entities to contract; *consensus ad idem* between such parties; intention to create a legal relationship, through promises that are enforceable by law; promises supported by consideration.[26] Going back to the philosophy behind AI, it is quite right to substantiate that there would a lack of intention and thus, this relationship falls short of being governed by Contract Law. Further, the aspect of consideration would not be fulfilled because the computer merely is following the instructions and that there is no benefit that it derives from such a relationship.

Nevertheless, in future, if it so happens that the Jurisprudence of Contract Law accepts such a relationship to be of the nature of an agency, then the liability so arising by the actions of the AI will entail that of vicarious liability. An agent will make the principal responsible so long as the agent does the act within the scope of his authority to do so or does so under the actual control of the principal.[27] The AI, if crudely put attempts to maximize the output to reach the goal. In case of the Google Duplex, the AI is directed to make appointments and reservations and it does so. Since an AI is not as autonomous as humans, it is highly possible that it features as the perfect agent, an extension of the principal himself. Thus, the principal is vicariously liable for all acts that are the resultant of the AI's actions.

In case of criminal liability, the principal is not held vicariously liable. But under circumstances as appearing in the fictional novel The Origin, by Dan Brown, it can be noticed that the AI is hardwired to reach its goal and would do absolutely anything to achieve it, even if it requires hiring contract killers. While the fiction is a stretch, it clearly highlights the fact that if unsupervised, the AI can go rogue. In such scenarios, the criminal liability should be accrued to the principal, due to whose negligence a crime would have been committed. Under such circumstances, the theory of strict liability is applicable.

The concept of Strict Liability first originated in the case of *Rylands v Fletcher*. The basis for liability was laid down by Blackburn, J. which was approved by the House of Lords. "The rule of law is that the person who, for his own purpose, brings on his land and collect and keeps there anything likely to do mischief if it

[26] Ted Allen & Robin Widdison, Can Computers make Contract? 9 Harv. J.L. & Tech. 26, 30 (1996)
[27] Sitaram Motilal Kalal v Santanuprasad Jaishankar Bhatt AIR 1966 SC 1697

escapes, must keep it at his peril; and if he does not do so is *prima facie* answerable for all the damage which is the natural consequences of its escape."[28]

In this analogy, the principle requires slight tweaking with respect to AI. The damage caused by natural consequences of escape can be a relayed as any undesirable, rogue activities as a consequence of the command that is given to it. Thus, if the AI's activities lead to an untoward result, then the principal attracts strict liability, as a result of negligence of not keeping the AI in control.

- **Producers' Liability with special reference to self-driven cars**

While the sections dealt above regard AI's personality, this particular section treats it as a product, analyzing it from the point of view of Consumer Law. The concept of Product Liability can be understood as the legal responsibility of sellers to compensate buyers for the losses suffered because of defects in the goods purchased.[29] Here, the AI machines and soft ware are considered as products and any injury, arising thereof shall be the liability of the producer. The producer (seller/manufacturer/distributor) attracts liability in case of negligence. Such negligence may arise from a duty to inform; improper design; failing to inspect properly or simply, based on the nature of the good, following the principle *res ipsa loquitor* (the thing speaks for itself).

The concept of Strict Liability figures its way under Consumer Law also. Strict Liability holds a seller liable without his fault on his part for injuries arising from the marketing of a product which is considered to be defective or unreasonably dangerous. Thus, if at any point, the AI tends to cause any 'injury' the seller is bound compensate the victim.

Self-driven cars or autonomous cars are making an appearance, increasingly. In March, 2018, a woman was killed in Arizona by an Uber, which was in autonomous mode.[30] In the event of damage caused by a fully autonomous vehicle, The owner/ possessor of the vehicle has no application of delictual liability. The damage to the third party, is a result of a bug in the control program.[31] Unlike regular motor vehicle accidents, the owner/ possessor has little to do with how the car functions. It would then mean that the producer/ manufacturer is liable.

Under the Estonian Law, any movable, including electricity and computer software, is deemed to be a product, even when the movable forms a part of another movable, or an immovable. It would then mean that the fully autonomous vehicle, as a whole, as well as the computer program that controls it is considered to

[28] Fletcher v Rylands (1886) LR 1 Ex 265, 279
[29] Cornelius W. Gilliam, Products Liability in the Automotive Industry, at 3 (1960)
[30] Robert L. Rabin, Uber Self-driving Cars, Liability and Regulation, SLS Blogs (last accessed on 16/09/2018 at 23:55 p.m) https://law.stanford.edu/2018/03/20/uber-self-driving-cars-liability-regulation/
[31] Taivo Liivak & Janno Lahe, Delictual Liability for Damage caused by fully autonomous vehicles: the Estonian Perspective, 12 MUJLT 49, 56

be a product. Thus, going by the same logic, both the manufacturer of the body of the car and the program developer are to be considered as manufacturers.[32] The trick here would be to determine the immediate cause of the accident, be it the glitch in the program or with any part which is to be considered as hardware and the liability is to be attributed respectivel

CONCLUSION

The primary aim of the paper was to determine the nature of Artificial Intelligence and also to prescribe modes of assessment of liability. As it is outlined in the third section of this paper, with present technological advancement, it is highly likely that an AI system has "Will" or consciousness. Even as an application of Kantian Theory, an AI is hardly human as it lacks the morals and mores. It is very unlikely to program these facets, as morals tend to be of personal in nature and replication is unwarranted. It is thus, concluded that AI is to not have a separate recognition, but be recognized with a natural person, preferably, the creator or the owner.

The paper in conclusion, has prescribed attributing liability in three forms, considering the status of the AI as a jural person and attributing the liability to the natural person in terms of Parental Liability, Vicarious Liability or Producers' Liability. Following such prescription, it is also imperative to provide suggestions to deal with a large scale acceptance of AI systems in the world.

Every sovereign nation can reasonably set up a committee to inquire into the varied types of AI and their integration into the society. This committee may consist of stalwarts in the field of Science and Technology, Law and Policy-making to seamlessly incorporate the growing sphere of Artificial Intelligence into human lives.

There may also be autonomous boards to look into large scale production of robots and autonomous machines and to license and regulate such manufacturers. A suggestive would also include a separate legislation to govern the relationship between human beings and autonomous machines.

It is quite possible in the future that AI may replace human being, just as industrial machines replaced manual workers, post the Industrial Revolution. It is then required that such replacement is so done justly, providing compensations if required.

As Social Engineers, it is the bounded duty of the fraternity of Lawyers to incorporate the continuous changes in the society, creating a secure environment to the human beings, and in future, robots.

[32] *Id* at 62