ARTIFICIAL INTELLIGENCE: THE LIABILITY PARADOX

Gyandeep Chaudhary[*]

Abstract

In this modern era, technology has evolved rapidly, and significant technological application of Artificial Intelligence (AI) has become an integral aspect of human nature, and a significant shift has taken place in human life. The issue of ascertaining liability, civil and criminal, for damages or losses resulting from its activities becomes a matter of priority for AI, which exercises control over itself in various degrees. The critical concern is that either national or international law does not consider AI as a subject of law, implying that AI can not be held directly responsible for its actions and damage caused subsequently. Moreover, the issue of liability assessment also presents a complex question of whether or not AI should be granted legal personality. Given the preceding, a question naturally arises: who is responsible for the damage caused by the actions of AI ?

I. Introduction

MOST OF us by now have heard of AI. It has been hyped up as the technology that will change the world in the coming decades. It promises self-driving cars, customer service chatbots, programs that can think and learn almost anything by themselves, and much more. But what is the term AI really referring to?

There is an easy way to define AI and then there is the long-winded more abstract way to define AI. It can be defined as a sub-field of computer science. The goal of this field is to figure out ways to program computers to learn on their own, which can be accomplished by simulating "intelligence" since this is how humans learn. This form of intelligence would be created artificially by a programmer hence the term "Artificial Intelligence".

AI could be defined as the use of technology through which human intelligence could be mimicked, following a detailed, expansive and exhaustive process of creating as well as applying algorithms, using a dynamic technological environment within computing.[1] In simple

[*]Ph.D Scholar, Indian Law Institute, New Delhi.
[1]OECD, *Artificial Intelligence in Society* (OECD Publishing, Paris 2019).

words, AI is an extremely advanced technology that makes a computer think as well as act like a human being, with the help of programming within bots.

One could also define it as any kind of man-made entity that has the ability to use working memory to utilize cognitive functions like abstract reasoning and logical deduction and able to learn new things autonomously[2]. This entity would also have the ability to form long term plans into the future using these cognitive abilities. Of course, this definition won't accurately describe AI until we actually reach the point where the programs we create possess real intelligence. Current AI is far behind this benchmark, most programs are only able to perform autonomously in a very narrow domain which limits their functionality.

Henceforth, artificial intelligence systems over the past one decade have gained a rapid momentum within this extremely tech-savvy world, with highly technical and sophisticated technologies used to develop ingenious, intelligent as well as intellectual AI systems. Therefore, that day is not far aware when these smart bots will start producing useful and spectacular inventions without really taking the help of human intelligence.

This ability of AI in producing and generating information, content, inventions, technology, *etc.*, has raised a big question concerning the challenges and problems that it can give rise to concerning the legal liability of AI. Therefore, existing laws in most countries will not be sufficient enough to deal with the liability concerns arising due to actions and decisions of an AI system.

Though the issue of AI regulation that establishes the liability for damages is still to be ascertained in India, it is a global phenomenon which needs to be solved. It is apparent in light of the impact of expedited globalization efforts; the problem of AI is not indigenous or limited to a territory and its traditional legal practices. The absence of a regulatory framework in the field of AI is an issue for the worldwide community, including common law and civil law states.

As of today, AI's operation is unregulated, and there is no specific legislation which effectively deals with it; therefore existing mechanisms and legislations are to be used to establish the liability of the AI for its actions and subsequent damages if any. Currently, established legal

[2] Jessica Peng, How Human is AI and Should AI Be Granted Rights?, *available at:*http://blogs.cuit.columbia.edu/jp3864/2018/12/04/how-human-is-ai-and-should-ai-be-granted-rights/ (last visited Aug. 1,2020), *see also,* OECD, *Artificial Intelligence in Society* (OECD Publishing , Paris, 2019).

norms provide for due compensation given in case of any damages caused by any unlawful act of the offender and the compensation is to be given by the offender or the person responsible for the actions of the offender. Therefore, in light of this legal preposition and keeping in mind that no legislation or regulation applies to AI, a severe question stems out as to who will assume the legal liability and provide for compensation in case of any damages caused due to actions of AI?

Therefore, this paper aims at analyzing,both national as well as international laws, to determine the liability of artificial intelligence, and whether existing liability principles as applied to humans could be applied to AI (considering these AI software's, and bots are ultimately created or developed by humans).

Henceforth, even though these questions at first may seem to be quite confusing still with the growing pace of AI inventions, it has become the need of the hour for us to draw specific legal solutions to these complex questions. And therefore, the researcher has further provided for an in-depth analysis of the Indian legal framework and how it is to be interpreted as to construe answers to these complex questions highlighted by the international community concerning the use of AI.

India is slowly yet steadily moving ahead in the AI sector with major companies like Apple and Salesforce acquiring Indian companies Tuplejump and MetaMind respectively for AI-powered technologies[3]. Not only this, but the increase in AI start-up in India has been massive, with increasing amounts of funds currently invested in research and development of the same. One of the most notable facts is that how an AI space, Sentient, received an investment amount of 143 Million USD in its initial years.[4] Therefore, there is no doubt in the fact that with such a massive increase in AI within the country, the scope of such AI innovations touching upon the day to day life of human beings is not very surprising, and so this raises the need for understanding the existing legal framework of India, to determine various liability and rule of law under the Indian legal framework.

II. Determination of liability of Artificial Intelligence

[3]Salesforce Acquires AI Startup MetaMind, *available at:*https://fortune.com/2016/04/04/salesforce-metamind-acquisition/ (last visited August 1,2020).

[4]The 10 Most Well-Funded Startups Developing Core Artificial Intelligence Tech, *available at*: https://www.cbinsights.com/research/most-well-funded-artificial-intelligence-companies/?utm_source=CB+Insights+Newsletter&utm_campaign=686c2ed68eTop_Research_Briefs_7_9_2016&utm_medium=email&utm_term=0_9dc0513989-686c2ed68e-87590629 (last visited July 29, 2020).

From the law we can derive our legal rights and duties. To follow the law is thus to perform duties and to be granted rights[5]. Legal personhood for AI is consequently a question whether AIs should have rights and duties in accordance with the law. The solution may be futuristic and progressive, yet a sufficient analysis should involve a brief discussion about legal personhood for AIs, as it would make the AIs accountable for their own actions.[6]

AI criminal liability requires legal personhood for the AIs, and would be similar to corporate criminal liability that some legal systems are recognising. Corporate criminal liability is deemed to be a fiction; a construed form of liability where the corporation is attributed with its employee's acts.[7] In contrast with corporations, AIs would be accountable for their own behaviour, not attributed with anyone else's. Even though it seems to be a simple solution not violating the rule of law requirements, it requires a more comprehensive assessment than this.

When a person or individual commits a crime against another person then he or she will definitely be subject to the criminal law which has been defined within that particular state. However, when it comes to artificial intelligence then any crime which is committed through it against mankind, may not be categorized as a conventional crime considering it has been committed using a software or a robot which is actually separate from the human who has invented that particular software or program or robot. Therefore in order to determine the criminal liability of acts which are committed through artificial intelligence, it is first important for us to understand whether or not artificial intelligence in itself is a legal entity and what are the key challenges which have taken place in determining and detecting the *actus reus and mens rea*, that is the act and mental (intention) element which is considered to be an essential measure for determining the commission of a crime .[8]

Is there a 'black box' problem?

Artificial Intelligence has become a large part of daily life for computer and smartphone users as they heavily rely on the complex problem-solving algorithms to perform even the most basic tasks efficiently and quickly. It is equally essential for these algorithms to function smoothly

[5]John Chipman Gray, *The Nature and the Sources of Law* (Cambridge University Press 1909) 27-28.
[6]Mireille Hildebrandt, 'Criminal Liability and "Smart" Environments' in R.A. Duff and Stuart P Green (eds) *Philosophical Foundations of Criminal Law* (OUP 2011) 506-32.
[7]Matilda Claussén-Karlsson, Artificial Intelligence and the External Element of the Crime: An Analysis of the Liability Problem, *available at*:https://www.diva-portal.org/smash/get/diva2:1115160/FULLTEXT01.pdf (last visited on 1 June 2020).
[8]Hallevy G., The Criminal Liability of Artificial Intelligence entities, *available at*: http://ssrn.com/ abstract=1564096 (last visited on April 15, 2020).

and flawlessly, and for us to understand how they function, which aids in further refinement of the algorithm. However, when we try to understand the functioning, we encounter a dead-end, and it becomes impossible to explain what is going inside the AI.

It is a significant issue, yet right now, it is constrained to gigantic deep learning models and neural systems. Since AI systems comprise complex algorithms and data sets that are software-generated rather than developed by human beings, these neural systems firstly break the problem in hand into zillions of pieces and afterward linearly process them bit by bit in order to realize a realistic output. Since the human mind does not work similarly; therefore, we do not have any mode of realizing what precisely the calculation the neural system is doing, or strategies are applied. Therefore this phenomenon is known as the "black box" or "explainability" issue as, throughout this problem-solving process, there is no method of obtaining access inside the neural network, which will give a view of ongoing processing.[9] This not just keeps us from gaining profound knowledge required to change the algorithm and subsequent calculations, but it causes a wide range of issues with trust involving AI systems or neural systems. Accordingly, it is said, it would likely never be possible for a human to describe why a sound AI system using such self-generated methods or data sets reached a specific answer or made a particular "decision."

We do think it is essential to bear in mind that any consideration of liability in a civil or criminal matter is ultimately a question of whether or not the acts or omissions of the relevant defendant as brought about by the applicable AI framework's choices and decisions were illegal. Did those acts or omissions amount to breaches of contract, negligence, or criminal offenses? It is essential to reinforce the point that the defendant, regardless, will be a legal person, not an AI framework.

To address these kinds of queries, a court will not have to comprehend why the pertinent AI framework settled on the choice that prompted the defendant's supposedly unlawful act or omission. We cannot help thinking that it is not feasible for us to determine why the AI framework settled on a choice. However, the AI framework did undoubtedly, in reality, settle on that particular decision, which either caused or added up to the defendant having committed an illegal act or omission.

[9]Niti Aayog, "National Strategy for Artificial Intelligence" 85-86 (June 2018).

For example, a plaintiff acting on a piece of poor investment advice given by the defendant generated by an AI system suffered a loss. Plaintiff may argue that the defendant failed to apply reasonable care and skill in his duty to the investment advice as per the implied terms of service contract. Without any explanation as to why the robo-advisor generated that poor advice, the plaintiff may be able to establish a breach of duty and the defendant's duty to take reasonable care. Therefore, it may be possible for courts to hold the defendant at fault based on the nature and quality of opinion given to the plaintiff.[10]

The court may have ruled in the same way irrespective of whether the advice produced using a robo-adviser or a human adviser. We should not forget that the human brain also has "black box" features (we often cannot explain human behavior), but this has not prevented courts from finding defendants liable in the past.[11]

It would be pretty early to conclude whether the "black box" will give rise to a legal problem while determining the legal liability of AI systems' decisions or not. However, in most cases, the inclusion of an AI system with a "black box" would hurt the plaintiff's ability to establish the defendant's act or omission, resulting in an unlawful act.

This paradox is because, under English law, a claimant will need to show:

i. Factual causation that the consequence would not have occurred but for the defendant's actions; and

ii. Legal causation, a complete chain of causation between the defendant's actions and the consequence (the defendant's act need not have been the sole cause of the consequence, but it must have made a significant (or more than minimal) contribution to that consequence).[12]

Neither of these tests requires the claimant to explain why the defendant acted in the way contended.

Where a defendant's state of mind is relevant to determining their liability, the law that applies will depend on the nature of the wrong alleged. However, it is imperative to take note that

[10]When AI systems cause harm: the application of civil and criminal liability, *available at*: https://digitalbusiness.law/2019/11/when-ai-systems-cause-harm-the-application-of-civil-and-criminal-liability/#page=1 (last visited on May 10,2020).
[11]*Ibid*.
[12]*Ibid.*

eventually, it will be pertinent to watch the respondent's perspective; additionally, the choice made by an AI framework would be likewise applicable. It may be possible and conceivable to establish that the defendant has the fundamental and essential perspective or viewpoint without explaining why the AI framework settled or made a specific choice.

Artificial intelligence, a legal entity or not

The first death that ever took place because of a robot within this world was that of Kenji Udhara, the engineer who worked in Kawasaki heavy industries plant wherein a robot was deployed to do specific manufacturing work. Therefore, when Kenji was repairing the robot, he forgot that he had to shut the robot because of which the robot detected Kenji as an obstacle after which he was brutally pushed towards an adjacent machine by the powerful hydraulic arm of the same robot, resulting in the death of Kenji almost instantly.

At that time and even presently, various laws of states around the world are unable to provide for any concrete criminal legal framework to deal with such instances officially wherein robots are involved is a commission of a specific crime or injury to an individual.

With the invention of AI , various significant dimensions have been added to the world and to deal with such a rapid pace of development. It is equally crucial for states to legislate which would bring more clarity upon the status of those instances and crimes which usually take place using robots for Artificial Intelligence Software.[13]

Legal status of artificial intelligence

Indian legal system does not explicitly have any legislation or statute which states about the unborn child and rights attributed to it. However, some statutes[14] not only recognize and state about an unborn child, but they also define such a child to be a legal person who acquires such rights only after taking birth. However, as a grey area in the legal domain, the legislature is silent on the notion of protection granted to such unborn and duties owed to such unborn which is inherently problematic. In the same way, AI systems are still at a very nascent stage and

[13]*Supra* note 8.

[14]The Transfer of Property Act, 1882, s. 13: which deals with the transfer of property for the benefit of unborn defines,"*Where, on a transfer of property, an interest therein is created for the benefit of a person not in existence at the date of the transfer, subject to a prior interest created by the same transfer, the interest created for the benefit of such person shall not take effect unless it extends to the whole of the remaining interest of the transferor in the property*."

Indian legal system still does not recognize it, which is an alarming state of affairs, let alone the attribution of any rights or duties and liabilities upon AI systems.

The legal status of a person or entity is directly linked to its autonomy because of which this status is not only granted to humans but cooperation, companies as well as organizations too. But when it comes to artificial intelligence, then it is not recognized yet by any legal system as legal entity except for Saudi Arabia wherein a robot called Sophia which is the state has recognized an artificially intelligent humanoid as a citizen of the country with rights and duties equivalent to that of human beings a noble person living within the state. The question of giving legal entities to artificial intelligence robots or software is subject to the fact that whether they can be entrusted with certain rights and duties with which a living person is usually entrusted upon.

While a living person is autonomous and he or she has the right to take his or her own decisions, on the other hand, an artificial intelligence system is created by humans and works as per the directions of the programs which have been added within its system to perform particular tasks and that too in a specific manner, but is capable of working autonomously.

Even though corporations or companies confer with the status of a separate legal entity, they are still equally liable to the stakeholders for any liability that may occur in the future from such transactions which have been entered into by these corporations or companies.

However, in case of artificial intelligence even though humans create it, it is still entirely independent and can perform such tasks which may be a consequence of a malfunction or wrong programming that can result in the commission of crimes even when the same is not intended on the part of the creator of such AI software.

The criminal liability of artificially intelligent robots is not clear under the State Law of any country. Thus, subjecting violations of laws and commission of crimes through artificial intelligence to only judicial pronouncements, serving as the primary source of the decision upon such cases where artificial intelligence is responsible for committing a specific crime (involving or excluding the directions of the creator which created such artificial intelligence robot software or algorithms).[15]

[15] *Supra* note 8.

The question of ascertaining liability, both civil and criminal, of an AI entity, parallelly impinges upon whether legal personhood may or may not be granted upon it. Kurki & Pietrzykowski[16] debate over the concept of legal personhood against the background of its moral and legal applications on the backdrop of normative jurisprudence. While analysing the connection between humanity, legal personhood and legal personality, they discuss how practical and several financial reasons could become an important factor for granting legal personhood to AI systems. They examine the concepts of "personism" and "personalism" and highlight as to how the personality may be separated from humanity.

The attribution of legal personhood has been addressed by Kelsen[17] in his theory of personality, according to which, granting of legal personhood is only a 'technical personification' to assert rights, duties and liabilities. The theory implies that legal personhood of an entity is, in general, a legal device to organise its rights and liabilities. Based on the Hohfeldian analysis of rights,[18] every right has a corresponding duty as its jural correlative. In the light of a jurisprudential analysis of these theories, the question of whether robot rights and liabilities may be rightly asserted by granting them legal personhood is examined. Granting legal personhood may, in turn, result in limited liability for the humans concerned with manufacturing or programming or operating the AI system. It is argued that, perhaps, at this point of time when the technology is still being developed and experimented in newer fields, granting of legal personhood to an AI entity for ascertaining liability may not be necessary in order to make it liable immediately. However, certainly in the near future with growing over-dependence on AI systems a need may arise to grant it personhood to determine the exact quantum of liability of such systems.

While the legal theory of artificial agents is not fully fleshed out, whatever path is adopted will have a significant effect on philosophical theorizing about artificial agents. The non-reliance of computing and mentality on a particular physical substrate has made possible speculation that the cognitive status of agents will be a matter of pragmatic judgement. The high point of such pragmatic deliberation is the legal sphere.[19]

Therefore, due to absence of legislative framework and specific policy guidelines vis-à-vis Artificial Intelligence systems in India, it is bound to attract several ethical and legal issues

[16]Visa A.J. Kurki,Tomasz Pietrzykowski (eds.) *Legal personhood: Animals, artificial intelligence and the unborn* (Springer Cham, 2017).
[17]Kelsen, H, *General Theory of Law and State* (Harvard University Press, Cambridge, MA,1945).
[18] Heidi M. Hurd, Michael S. Moore, *The Hohfeldian Analysis of Rights* 63 *TAJ*J 295 (2018).
[19]Samir Chopra, Laurence White, *Artificial Agents - Personhood in Law and Philosophy* 635-639 (2004).

upon its usage or application. Hence, the requirement of policy guidelines for corporations (creator, developer, manufacturer, and software programmers of AI systems) and the legislature to meet various ethical and legal standards, could primarily be addressed by determining the nature of AI systems as an entity. Therefore, accordingly the liability may or may not be shifted from creators to the AI system which exercises some degree of self-control.[20]

Criminal liability

Renowned legal researcher and lawyer Gabriel Hallevy proposed that certain AI systems can meet the essential requirements of criminal liability,

i. Constituting *actus reus,* an act or omission ;
ii. Moreover, *mens rea*,requiring knowledge or information, and
iii. Strict liability offences, where *mens rea* is not required.

The requirement of *actus reus* in proving criminal liability

The criminal liability of artificial intelligence robots and software whenever a criminal act is committed, then the general basis for the same is the requirement of law. Therefore without an *actus reus*, the criminal liability of an individual cannot be proved, and so in the case of artificial intelligence, *actus reus* can only be established if the crime which committed through such a mechanism can be prescribed to a human being so that the very condition of commission of an act can is satisfied to punish and prove the criminal liability of an individual[21].

The element of *mens rea*:

When it comes to *mens rea*, the prosecutor is required to prove, an act that was committed by an AI was intentional on the part of the user against another person. The highest level of *mens rea* is knowledge which may or may not be backed with the intention of a particular user under whose supervision or administration a specific act was committed by an artificial intelligence robot.

[20] Priyanka Majumdar, Bindu Ronald *et.al.,*"Artificial Intelligence, Legal Personhood and Determination of Criminal Liability"6 *Journal of Critical Reviews* 323 (2019).
[21] Legal liability issues and regulation of Artificial Intelligence (AI) Dissertation work -Post Graduate Diploma in Cyber Laws and Cyber Forensics Course Code: PGDCLCF Submitted by: Jomon P Jose, *available at*: https://www.academia.edu/38665852/Legal_liability_issues_and_regulation_of_Artificial_Intelligence_AI_Dissertation_work_Post_Graduate_Diploma_in_Cyber_Laws_and_Cyber_Forensics_Course_Code_PGDCLCF_Submitted_by_Jomon_P_Jose (last visited on June 15, 2020).

The lowest level of *mens rea* is when criminal negligence or recklessness is present on the part of the user of such an AI machine, which would have been otherwise known by a reasonable person under his strict liability.

Hallevy proposed three legal models to be considered to examine offences committed by AI systems:[22]

i. The Perpetration by Another Liability Of AI

ii. The Natural Probable Consequence Liability Of AI

iii. The Direct Liability Of AI

i. **Perpetrator-via-another**, in case of commission of an offence by any mentally challenged person, a minor or an animal, in that case, perpetrators are an innocent agent due to fact that they lacked the requisite mental capacity to constitute a *mens rea*, and this is applicable in case of strict liability offences too. However, if that innocent agent acted on someone else's instructions, then, in this case, the person giving the instruction or the instructor would be criminally liable, for example, a person training his dog to attack strangers in case of any specific event.

Therefore, as per this model, while considering AI systems or programs as an innocent agent, the user or the system developer could be identified as perpetrator-via-another.

ii. **Natural-probable-consequence**, in this model, part of the AI program envisioned for bonafide purposes is wrongly triggered, which leads in the performance of a criminal act. Hallevy gave an example of Kenji Udhara, the engineer who worked in Kawasaki heavy industries plant wherein a robot was deployed to do specific manufacturing work. Therefore, when Kenji was repairing the robot, he forgot that he had to shut the robot because of which the robot detected Kenji as an obstacle or threat to its assigned duties and calculated that the most efficient way to eliminate this threat was by pushing him into an adjacent operating machine. He was brutally pushed towards an adjacent machine by the powerful hydraulic arm of the same robot, resulting in the death of Kenji almost instantly and then resumed its duties.[23]

[22]The Basic Models of Criminal Liability of AI Systems and Outer Circles, *available at:* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402527 (last visited on June 15, 2020).
[23]Yueh-Hsuan Weng,Chien-Hsun Chen *et al.,* "Towards the Human-Robot Co-Existence Society: On Safety Intelligence for Next Generation Robots" 1 *Int J Soc Robo*. 273 (2009).

This model is used to establish "natural or probable consequence" liability or well known as "abetment" under Chapter V of The Indian Penal Code,1860 (*hereinafter* referred to as IPC), which governs the liability of persons considered as an abettor in the commission of an offence. Hallevy discusses United States Criminal law where an accomplice could be found liable for the act even if no conspiracy is proved, provided the act of the perpetrator was a natural or probable consequence which the accomplice encouraged or aided and had the knowledge[24] that a criminal conspiracy is underway.

In Indian criminal law, section 111 of IPC under Chapter V provides for the principle of probable consequence, which provides that in case of act abetted and act done are different. The abettor will be liable for the act done by the perpetrator in the same way and to the same extent as if he had directly abetted it, with the only requirement of a possible consequence of abetment. The general notion regarding abatement is that there could be no conviction for abetment unless there is a commission of an act. However, in some instances where the proof would be insufficient to charge the perpetrator but sufficient to convict the abettor, the perpetrator could be acquitted, and the abettor is likely to be convicted based on evidence and facts.

Therefore, AI system developers and users may be held liable for the act of the AI system if they knew that the act done is a natural or probable consequence of the usage of their AI system. However, while applying this principle, a distinction must be made between the AI systems explicitly developed for criminal purposes and those with legitimate other purposes, *i.e.,* where the AI system knows about criminal intention and where it does not know. This model covers the former group of AI systems while the prosecution of the latter group may not be possible due to a lack of knowledge (but strict liability would apply to them).

iii.**Direct liability**, this model attributes both *actus reus* and *mens rea* to an AI system.

It is comparatively easy to ascribe an *actus reus* to an AI system. If by chance the outcome of any action taken by an AI system ends up being a criminal act or its failure to act in a situation where there was a duty to act, at that point the *actus reus* of that offence has happened. Attributing out a *mens rea* is very hard, thus it is here that the three-level of *mens rea* becomes substantial. Whereas, in case of strict liability offences where intention need not be proven or

[24]Francis Bowes Sayre, "Criminal Responsibility for the Acts of Another"43 *Harv. L. Rev*. 689 (1930).

is not required, it is maybe possible to hold an AI system liable for the criminal act. Strict liability may be understood by an example involving an autonomous self-driving car and overspeeding, where overspeeding is a strict liability offence. So, as per Hallevy's model, then the law regulating the criminal liability of over-speeding could be possibly applied in the very same fashion as applied to humans on an AI program which was driving the car.

Civil liability

When software is defective, or when a party is injured as a result of using software, the resulting legal proceedings normally allege the tort of negligence rather than criminal liability.[25] Gerstner[26] discusses the three elements that must normally be demonstrated for a negligence claim to prevail:

i. The defendant had a duty of care,

ii. The defendant breached that duty,

iii. That breach caused an injury to the plaintiff.

In case of defendant's duty of care, Gerstner points out that undoubtedly there is the duty of care attributed to the software or system vendor toward the customer; however, it is not easy to determine what quantum of standard care is required. If the said system is an "expert system", then the level of standard of care would be of a professional at least if cannot be of expert level.

In case the defendant breaches a duty, numerous ways suggested by Gerstner wherein breach of duty of care by an AI system could take place such as,

- developer's failure to detect errors in program features and functions,
- an inappropriate or insufficient knowledge base,
- inappropriate or insufficient documentation and notices,
- failure to maintain an up to date knowledge base,
- error due to user's faulty input,
- excessive reliance of the user on the output,

[25]Tuthill G.S, Legal Liabilities and Expert Systems, *AI Expert* 1991.
[26]Gerstner M.E, Comment, Liability Issues with Artificial Intelligence Software 33 *Santa Clara L. Rev.* 239

- misusing the program.

Lastly, in case of injury to the plaintiff due to breach, can AI systems cause or suppose to cause an injury is debatable. However, the critical question involving AI is whether AI systems suggest any solution in a particular scenario just like most of the expert systems or instead AI system itself takes the decision and acts accordingly, for example, an autonomous car. Therefore, while the former case involves at least one external agent which makes it difficult to prove causation wherein in latter due to no involvement of an external agent proving causation is comparatively easy.

III. Policy Regime

In the year 1996, prominent researchers Tom Allen and Robin Widdison reasoned "soon, our autonomous computers will be programmed to roam the Internet, seeking out new trading partners e whether human or machine… At this point we must inquire whether existing contract law doctrine can cope with the new technology, and if so, how."[27] They concluded, "neither American nor English law, as they currently stand, would confer legal status on all computer-generated agreements."[28] It implies that the legal doctrines in existence at that time could not adapt to the damage done by technology. Furthermore, it led to the emergence of an issue: to decide how the existing law should be changed.

Almost two decades have elapsed since Allen and Widdison published, and the contract done through the interaction of interactive voice response systems (IVRS) is now recognized and legally binding,[29] the simple question still looms at large: whether the existing legal doctrines can deal with the new, emerging and sophisticated technologies and the damage made by AI, and if so, how?

UNCITRAL's explanatory note features a general guideline revered in article 12,[30] which states any computer or machine programmed for a person (natural person or a

[27] Tom Allen and Robin Widdison, "Can computers make contracts?" 9(1) *HJLT* 28-29 (1996).
[28] *Id.* at 52.
[29] Art. 12, of United Nations Convention on the Use of Electronic Communications in International Contracts, "A contract formed by the interaction of an automated message system and a natural person, or by the interaction of automated message systems, shall not be denied validity or enforceability on the sole ground that no natural person reviewed or intervened in each of the individual actions carried out by the automated message systems or the resulting contract."
[30] *Ibid.*

legal entity) would be liable for any generated message by that computer or machine.[31] The Electronic Communications Convention Explanatory note section (213) of Article 12 outlines that: [32]

> Article 12 is an enabling provision and should not be misinterpreted as allowing for an automated message system or a computer to be made the subject of rights and obligations. Electronic communications that are generated automatically by message systems or computers without direct human intervention should be regarded as 'originating' from the legal entity on behalf of which the message system or computer is operated. Questions relevant to agency that might arise in that context are to be settled under rules outside the Convention.

Under Indian legislative regime no specific or separate provision provides under any legislation concerning the liability of such acts, which may be committed by a user, administrator, or producer through artificial intelligence software or systems. In India, the strict liability doctrine corresponding to the direct liability model proposed by Hallevy, is not as developed despite being a common law system as compared to English law. UK's criminal law, strict liability doctrine has evolved over the period time by accumulating existing English laws with modified, amended provisions, authoritative decisions of the judiciary, and statutory enactments made by parliament from time to time. On the contrary, an exhaustive codification of Indian penal laws leaves no scope for the judiciary to go beyond existing statutes.

If we examine the IPC then, chapter IV (general exception) mainly deals with matters of the existence of which negate the existence of such an intent. The definition of offences generally contains a reference to malicious intent to exclude all acts where such an intent is not present. Even where the definition is silent regarding the intent, it is held that on general principles, a malicious intent must be imported into the definitions of all strictly criminal offences.

In India, offences defined in IPC are careful to include *mens rea* in the definition itself, and the chapter of general exceptions very exhaustively. The definitions in the Indian penal code, along with the chapter of general exceptions are perhaps sufficient to exclude all cases to which a *mens rea* cannot be attributed.

---

[31] Ugo Pagallo, *The Laws of Robots: Crimes, Contracts, And Torts* 98 (Springer 2013).

[32] Paulius Cerka, JurgitaGrigiene, *et.al.*, "Liability for damages caused by artificial intelligence" 30 *CLSR*8 (2015).

On analysing of the definition of offences in IPC, they generally comprise the following principal elements:

i. A human being,[33]
ii. An intention on the part of such a human being to cause a particular consequence considered injurious to individuals or society, a malicious intent,
iii. The act being done,
iv. The resultant consequence.

However, there are a few cases where words indicating intention are missing from the definition of an offence. In such cases, either consequence of the act is so harmful to the state or society that irrespective of any intention it is deemed just and expedient to punish them, or cases where the actions are of such nature that they raise a strong presumption that the actor must have intended to do it. Section 121 (Waging, or attempting to wage war, or abetting waging of war, against the Government of India), 124A (Sedition), 359-363 (Kidnapping and Abduction) are examples of former, while section 232 (Counterfeiting Indian coin) is an example of later. Nevertheless, in the case of AI systems determining the *mens rea* would be humongous owing to the Black-Box problem, but the doctrine followed in the above scenarios might be applicable to determine the direct liability of the AI system.

The principle of probable cause liability or abetment under the IPC is sufficient to determine the offence and the penalty for abettors. However, at the given pace of development of technology, the legislature has made attempts to plug the gap and give more realm to it, and the Information Technology (Amendment) Act, 2008, widened the meaning of abetment to include act or omission by use of encryption or any electronic method.

Information Technology Act, 2000(*hereinafter* referred to as IT Act), which tries regulates all the aspects of modern day technology tries to define computer and related terms such as software etc, but Internet of things, data analytics as well as AI all of these three aspects are not covered under the IT Act, and neither are the liabilities which that may be committed by humans using these IT mediums (specifically AI software).Considering the primary intention of the Act was to provide a digital signature and electronic records a legal status, the Indian

[33]*Supra* note 5., s.10 defines *"Man", "Woman".—The word "man" denotes a male human being of any age; the word "woman" denotes a female human being of any age*., s.11, which states *"Person".—The word "person" includes any Company or Association or body of persons, whether incorporated or not.*

legislature really did not emphasize the scope of liability arising out of actions of AI and measures to combat the same.

Liability for damages

Under the Indian legal framework, there is no specific Indian legislation which deals with the criminal or civil liability of such crimes committed using AI. Therefore, it is to be noted that India is one such country which is moving towards the implementation of such policies through which we can incorporate AI within the entire government system, but at the same time, the legal system is ignoring the possible adverse impacts of cybercrimes that may be committed in the future, by using these highly technological and advanced AI systems.[34]

Role of judiciary in deciding the liability

In the legislative vacuum, the definition of final punishment as well as the criminal/civil liability of such acts which are committed by AI systems, software, and robots against other individuals, the only ray of hope which is left within the Indian legal system to tackle such cases is the Indian judiciary.

Even though there haven't been any significant landmark judgment which can provide for a breakthrough upon the guidelines of the use of artificial intelligence software or robots, to prevent the commission of any criminal or civil offense against others, but it is expected of the judiciary, with increasing pace of development through artificial intelligence to pass such guidelines as well as judicial precedents through which the use of artificial intelligence could be dealt with effectively by defining the criminal and civil liability of such artificial intelligence systems which may cause harm or damage to other individuals through various unethical practices such as phishing, hacking and data theft[35]*etc*.

IV. Conclusion

A future with intelligent AI-powered robots and technology may seem daunting at first, but I believe our future is bright, and the possibilities stretch far beyond what we are capable of comprehending right now. Mostly what I have seen lately is experts describing the risks of AI and painting a picture in which a terminator like doomsday scenario takes place instead of

[34]Indian Law is Yet to Transition into the Age of Artificial *Intelligence, available at: https://thewire.in/law/indian-law-is-yet-to-transition-into-the-age-of-artificial-intelligence* (last visited on March 18, 2020).
[35]*Ibid.*

talking about the possible benefits and how we can use this technology to better ourselves, create an ideal world, and even to explore other worlds. This pessimistic mindset is not helpful, and I believe we should not allow this to discourage progress in the field of AI.

The AI industry as a whole has been moving very fast. Companies tend to be quick to adopt new technologies out of fear of being left behind. Machine learning and deep learning allow companies to recognize patterns by analyzing increasingly large data sets which opens up bold new possibilities. With these new possibilities come a whole host of new ethical problems including but not limited to:

i. Legal issues arising out of the liability paradox,
ii. Intellectual property rights concerns involving advanced AI programs capable of self-generating content.
iii. Privacy concerns over the use of personal data,
iv. Discrimination by AI programs playing a role in the hiring of applicants,
v. Facial recognition,
vi. The use of autonomous military weapons which allow AI to make its own decisions regarding when to kill someone,
vii. Self-driving vehicles which will have to decide on their own what to crash into should they malfunction.

Today AI is not being recognized as a legal person both in National and International laws, which undoubtedly corresponds to it not being able to hold accountable for any damages caused by it. Therefore the principle enshrined in article 12 of the United Nations Convention on the Use of Electronic Communications in International Contracts maybe applied concerning the liability of AI, which states that the person at whose behest the system was programmed should eventually be held liable for any act done or message generated by that system

In light of the above proposition, the direct liability model as proposed by Hallevy can be applied, which translates that strict liability principle could govern the behaviour of the AI system, but it would leave other actors out of liability. Therefore a new concept of AI-as-Tool could be evolved and applied wherein taking the strict liability from Hallevy's Direct Liability model could be applied to an external agent(natural or legal person) who instructed the machine to act deriving from Hallevy's Perpetrator-via-another model, irrespective of the fact that whether such action by an AI system was planned and envisaged or not.

So when an AI system treated as AI-as-Tool, strict or vicarious liability could be easily applied to the damages done by an AI system. However, how an AI system operates and it's operating principles, i.e. independent decision making would make it challenging to establish the burden of proof appropriately. It would be an immense task for the plaintiff to prove the fact that there was a defect in the AI system when supplied by the original equipment manufacturer since AI is a self-learning system it is practically beyond human capability to distinguish between damage caused due to a product defect or resulting from the act done by AI system during its processing.

Therefore, practical applicability of such a liability model solely depends on the intent of the legislation as to how to tweak the existing law or enact new legislation to deal with the liability of AI systems in an era when the dominance of AI in human lives is increasing day-by-day.

Most governing bodies seem to be burying their head in the sand instead of thinking about the future of AI. Business owners seem to be more on top of the need for regulation because they have a clearer understanding of its potential. Countries like the United States are probably concerned about falling behind countries like China that are pouring billions of dollars into AI research and development. However, it is clear that when it comes to ethics and regulations regarding AI, China is a shining example of what not to do. We need to have our priorities in line moving forward.