Journal of Advanced Zoology

ISSN: 0253-7214
Volume ***44*** *Issue* ***S-5 Year 2023*** *Page* ***1478:1483***

Cybersecurity In The Age Of Artificial Intelligence – Who Should Be Culpable?

Miss Astha Chaturvedi[1*], Ruchi Tiwari[2]

[1]Research Scholar, Parul Institute of Law, Parul University*
[2]Associate Professor, Parul Institute of Law, Parul University

**Corresponding Author's E-mail: Miss Astha Chaturvedi*

Article History

Received: 01 June 2023
Revised: 05 September 2023
Accepted: 27 October 2023

CC License
CC-BY-NC-SA 4.0

Abstract

The age of the AI has come in hurtling and it has become difficult for law makers to keep pace. At the same time, unbridled by legislations, criminals in the cyber space are using these Artificial Intelligence tools. The instances of ChatGPT's use in writing academic articles to morphing pictures of people, the degree of seriousness is high. Another incident was the use of morphed pictures of a celebrity couple's child for disseminating fake news. The lines of reality and morphed is blurring to an extent that the line is very unclear and hence it is essential that proper laws are put in place to identify and assign culpability to the use of AI for criminal activities. Although the AI applications are guided by moral codes, the said moral codes would be guided by the developer's morals and hence the question of standardized laws, ethics and morals would arise here. While the debate around AI and intellectual property rights have been on the round for some time and Courts in certain jurisdictions have already deliberated on it, the question has shifted to criminal culpability in case of hacking, creation of photos that falsify evidence, publication of pictures of individuals that have been digitally altered via AI, etc. The question that arises is who is in control of the actions of the AI, since the creator of the AI is not always the person who is involved in criminal activities that are affected through the AI. Also, would a standard code of ethics be able to govern the AI's functioning successfully?

Keywords: *Cybersecurity, Artificial, Intelligence*

1. Introduction

Most dystopian and utopian worlds imagined in science-fiction have an element of Artificial Intelligence that led to the creation of such a world. The earliest account of a machine that possesses human-like intelligence is in the work of Samuel Butler's Erewhon, which was first published in the year 1872. However, the term that we are familiar with today, Artificial Intelligence (AI), was first coined only in 1956 by a John McCarthy as was state by Lewis, A Brief History of Artificial Intelligence.[1] The understanding of what constitutes an AI varies amongst the common man with a limited scientific understanding, but most of these definitions are justified since AIs differ according to their functionality. Nonetheless, artificial intelligence has been talked about and has been in ChatGPT and its widespread user interaction has made the arrival of Artificial Intelligence (AI) louder than ever. And notoriously it has been put to wrongful activities, like the instances of it being put to use by students to write essays, and thereby strengthening the argument of naysayers that AI is evil incarnate. OpenAI is being used to create realistic morphed photographs and videos that are difficult to distinguish from a real photograph. There are software applications that allow the alteration of photographs and videos that allow removal of background items etc. Although digital forensics would be able to determine if there has been tempering of the digital photographs, yet a common man would rarely make further investigations. Thereby, it may lead to tarnishing a person's reputation or at times mislead the public, thereby causing unrest in the society. An example of social media as a means of spreading fake news and posting of objectionable content can be that of the use of the popular software application, WhatsApp, which allowed the dissemination of messages without verification turned out to be the reason behind certain events. that took place across India. Thereby, a notification was released stating that since the application's groups allowed the dissemination of these messages faster, the administrator of the group would be held liable for allowing the distribution of unverified news.[1] The Bombay High Court deliberated on whether such criminal liability is acceptable and held that the guidelines could not be upheld[1]. In another case, that was heard in New Delhi, the High Court denied the assumption that an administrator of a WhatsApp group would be vicariously liable for

any defamation and loss of business caused to a person.[1] Although the judgement is sound, taking into account criminal jurisprudence, the answer to the question of culpability and justice still evades. In case of AIs, a second layer of problem gets added to the question of culpability, responsibility and answerability since a "machine" gets tied into the equation. The answer might lie in the determination of personhood of the "machine" or AI. The idea of personhood or juristic personality has been toyed with by various legal jurisdictions pertaining to topics of patent, copyright, product liability, discrimination and other issues.

2. Literature Review

A discussion about Ais usually conjure images of what fiction has visualized AI. However, AI comes in myriads of forms and with varying degrees of competence. This ranges from image recognition software that can be found in phones, to robots used in industrial manufacturing, Alexa and Siri as we know it – voice recognition AI, chatbots that can be found on almost every webpage and others.[1] The article by Daniel Barksy and other authors highlights various issues pertaining to AI. However, the initial classification of the AIs for understanding of the layman is noteworthy.[2] The biggest conundrum that exists regarding the development and use of AI is what has been envisioned by novelists imagining a dystopian world in which humans have to survive according to the human-like machines. The ethics and morals that would govern an AI is a factor that would influence the outcome of how an AI behaves.

The guiding ethics and moral codes can determine the biases a machine has, and hence it is an important element that needs to be considered while trying to bring in legislations governing AI and its use.[1] Gordon addresses how the evolution and rapid development of AI has an impact on a person's fundamental rights and further examines if the present-day AIs can be consider a person. His answer is a no, but his analysis can be used to develop on the onus of liability in case of misuse of AI. While we talk of morality and ethics, we work under the assumption that a machine is aware of its actions and its consequences just as a human being, who would be held liable for his actions if *mens rea* and *actus re* can be established. However, Andreas Matthias has identified, what he terms as 'responsibility gap' and this gap[1] There exist four potential gaps in accountability and each of these gaps caused by AI might is significant.

It has been suggested that a critical evaluation of incomplete and inadequate initiatives to address the responsibility gap: those who present it as a brand-new, insurmountable problem ("fatalism"), those who dismiss it as a false problem ("deflationism"), and those who confine it to just one of its dimensions or sources and/or who present it as a problem that can be easily fixed by the addition of new technical and/or legal tools ("solutionism"). A more comprehensive strategy to tackle this gap would be based on the notion of developing sociotechnical systems for "meaningful human control," that is, systems in line with the necessary human motivations and capabilities, to solve the responsibility gaps with AI as a whole.[1] Of these four, the research has concentrated on two gaps that are most relevant to the research. Attaching a juristic personality to a non-human entity gives it the capability to sue and be sued. However, as has been the case with companies, the lifting of the corporate veil becomes impertinent under certain circumstances (*Solomon v Solomon)*. Autonomy of an entity is a major factor in establishing its person-hood, but while companies have shareholders who can be held accountable for certain outcomes, the same is missing in case of AI. There would also be a lack of clear liability on the developer of the AI since any act that has been undertaken might not have directly resulted from the data or moral set that the developer had provided, but rather was an outcome of the learnings that the AI had gathered over the course of time.[6] AI is burdened with biases, and these biases also can contribute to how an AI act in a particular situation just as personal prejudices and ideas of morality effects the actions of a natural person. While the understanding is that the data sets provided to AIs are the cause of the bias that has been observed, National Institute of Standards and Technology (NIST) has a different tale to tell and insists that while we blame the AI biases on computational and systemic biases, researches often tend to overlook the human and systemic bias that are in place, which also forms a part of the data set received by AIs.

3. Materials And Methods

A major part of this research is theoretical in nature, based on ideas that have been worked on and debated by learned academicians in the field of AI and law. A large portion of the research is based on secondary sources, of which majority would be articles that have propounded theories of responsibility. The idea is to concentrate on a more specific area of AI and its potential of being an accomplice to morphing photographs and videos, and thereby causing liability for defamation, inciting violence, spreading public disorder etc. For this, a general study of criminal jurisprudence and the definitions under the Indian Penal Code and other relevant statutes have been taken into account and form a major part of the primary sources of the research. Furthermore, cases pertaining to recognition and derecognition of AI's personhood has also been delved into by taking into consideration case laws that have discussed on this topic.

However, a principal limitation of research on responsibility and culpability of AI is the lack of empirical studies, and hence the present research has made an attempt to develop an empirical case study to determine the importance of legislation on AI generated images and videos to illustrate how a common man is ill-equipped to distinguish a real photograph or video from that of a morphed one, and is further unaware of the potential

threats unverified pictures and videos pose. The sample size is small and is limited to educated subjects who have a minimum graduate level qualification, but has factored in people of various age groups. So, it would be able to give an insight into whether the lack of verification at the end of recipients of fake data and the need to address the onus of such data being spread. The data has been collected via Google Forms, and data has been analyzed through the data collected through it. The questionnaire and its responses form part of the research.

Moral Accountability

Morality is guided by various factors that a person grows up in and is molded into by the kind of household the person grew up in, religious belief, social circle, education and other influential factors. However, when the discussion is about an AI, how can its morality and moral accountability be determined? The logical answer would be that the developer of the AI would integrate the moral principles guiding it. As a person's environment effects its standard of morality, therefore the behaviour and decisions of the AI would also be either acceptable or unacceptable based on the State. Critical aspects of AI implementation have already made inroads in popular literature and culture. For instance, there are claims that one major weakness of pop-culture techniques is that they rely on proxies for identifying trends, such as a person's postal code or language in connection to their ability to handle a job or repay a loan, respectively.[1]

These correlations could, however, be unlawful if not discriminatory. Demographic discriminations were witnessed in AIs that used facial recognition system in software companies while undergoing a study. Therefore, the morality behind decision making suffers from prejudices, and these prejudices may exist because of implicit biases of the developer of the AI, sampling bias, existence of edge cases etc.[2] One of the four gaps of AI is morality accountability gap and although not as serious as culpability gap, it has a significant in the decision-making process. If an individual is asked to justify an action that they have undertaken, the individual feels an obligation to be more morally responsible in the decision-making process. In the philosophical literature on moral responsibility, moral accountability has been given a crucial position as an element for the justification and comprehension of moral responsibility practices.[2] Additionally, it aids people in connecting the world's events to their logical faculties, enhancing their sense of agency and accountability. Therefore, to hold an AI morally accountable, we must also bestow personhood to the AI.

However, morality has its own shortcomings, therefore a more favorable approach would be that AIs are developed with a standardized ethical code, since ethics can be made common to all unlike morality. The United Nations Educational, Scientific and Cultural Organization (UNESCO) has recognized that AIs, without ethical guidelines would add to real world problems like discrimination, and therefore have put forward certain recommendations pertaining to ethics for AI.[2] The guidelines provides that the AI should be aligned with principles of 'proportionality and do no harm', meaning that at all times it is to be ensured that the AI used is necessary to achieve and action and that such use will not interfere or be *ultra vires* to fundamental human rights; safety and security of human-kind in general is to be ensured; AI should not discriminate on the basis of age, culture, gender or any other such factors, The guideline further states that all stages of the AI life-cycle, a human or recognized legal entity should be made legally responsible for its actions. This principle clarifies that at all times the onus is to be on a recognizable person, or legal personal who can be held accountable for the acts of the AI.

Personhood, Culpability Gap and Legislation

The science-fiction movie *Robot & Frank* written by Christopher Ford and directed by Jake Schreier raised the question of criminal liability of an AI that was responsible for caregiving and its owner was a person who was suffering from dementia. Though it is a comedy, it begged to question that what if in the future it was no longer a fiction but reality. We have arrived at a time, a decade later, wherein liability for actions of an AI have become very real. The research has constrained its research to the question of liability of defamation and causing public disorder.

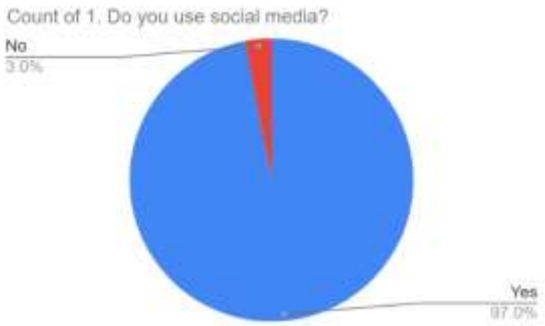
Section 499 of the Indian Penal Code provides that "*whoever by words either spoken or intended to be read, or by signs or by visible representations, makes or publishes any imputation concerning any person intending to harm, or knowing or having reason to believe that such imputation will harm, the reputation of such person, is said, except in the cases hereinafter excepted, to defame that person*", wherein it is evident that the key element to establish liability for defamation would be the *mens rea.* The establishment of intention to cause harm or the knowledge of an action being right or wrong would be difficult if not impossible to establish. The AI is unaware of moral and ethical codes of the society and to a large extent its morality is guided by its creator. It is one of the gaps of AI.[1] Moreover, the AI that has been generated is not necessarily used by the same person, an illustration of which would be driverless cars. Herein, unlike a company there is no clear accountable person behind the veil who can be held responsible for the acts of the person, and therefore attributing personhood would not be sufficient to establish culpability in criminal cases to serve justice. The example of an artificial intelligence assassinating moral responsibility as a result of the opacity and the complexity of AI may be that of a doctor using an AI driven system for diagnosing. The systems are based on the techniques of Deep Learning. This dataset requires detailed training on the nature of which is well known and clear.

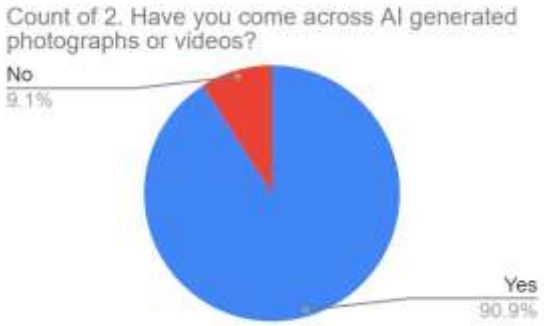
Furthermore, Sections 153A & 153B defines the offence of promotion of enmity between different groups of people who have differences based on religion, caste, language etc. Consider a situation wherein an AI generates a picture where a person can be identified by their clothing as to which religious denomination they belong to and showcases the harassment of another person belonging to another religion, which again is attributed to the clothing the person has on. This picture gets circulated in social media platforms like Twitter, WhatsApp, Snapchat, Facebook etc. In such a case culpability is hard to assume and a small empirical study showcases that the general public can be manipulated by AI generated pictures and videos.

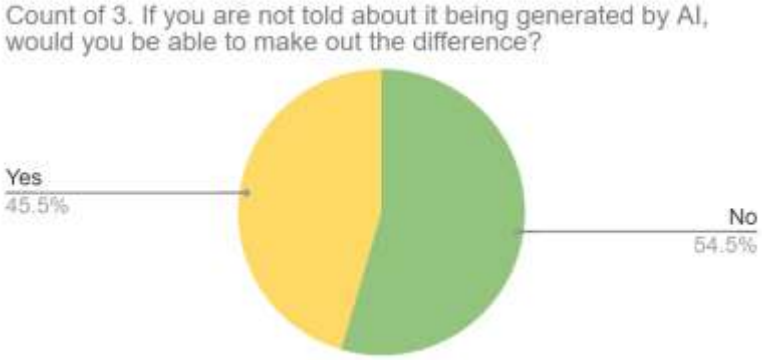
It is clear that a large portion of the sample does use social media and as much as 90% of the sample stated that they have come across AI generated photographs and videos. As much as 54.5% of which could not differentiate between an AI generated and a normally captured photograph/ video. It is necessary to reiterate here that the sample contains people from a minimum of graduate education level. Therefore, the likelihood of percentage of people who would not be able to differentiate or even realise that AI can be used for such work would likely be much higher. Since it is a educated sample group, 57.6% of the sample has opined that they would be inclined to conduct research on its authenticity. Interestingly, 3% of the sample stated that they would at times verify its authenticity while at other times, they might not take steps to verify the same. This would likely be the case when there is paucity of time to verify, or there might be other factors as well. The lack of insight into such factors also forms part of limitation of this study.

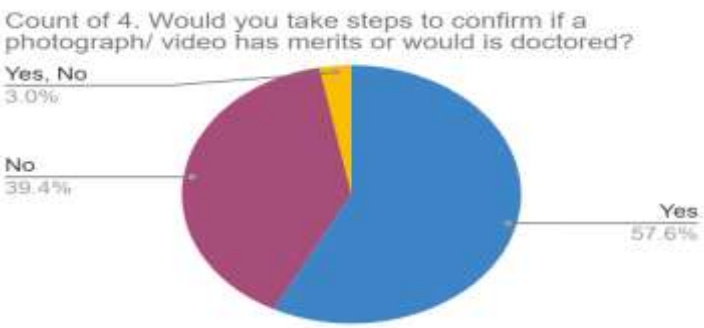
1. Do you use social media?	2. Have you come across AI generated photographs or videos?	3. If you are not told about it being generated by AI, would you be able to make out the difference?	4. Would you take steps to confirm if a photograph/ video has merits or would is doctored?	Gender
Yes	Yes	No	Yes	Female
Yes	Yes	No	No	Male
Yes	Yes	Yes	Yes	Female
Yes	Yes	Yes	No	Female
Yes	Yes	Yes	Yes	Male
Yes	Yes	No	No	Male
Yes	Yes	Yes	Yes	Female
Yes	No	No	Yes	Male
Yes	Yes	Yes	Yes, No	Female
Yes	Yes	Yes	Yes	Female
Yes	Yes	No	Yes	Female
Yes	Yes	No	No	Female
Yes	Yes	Yes	Yes	Female
Yes	Yes	Yes	No	Female
Yes	Yes	No	No	Female
Yes	Yes	Yes	Yes	Female
Yes	Yes	No	Yes	Female
Yes	Yes	No	Yes	Male
Yes	Yes	No	No	Male
Yes	No	No	No	Female
Yes	Yes	Yes	Yes	Male
No	Yes	Yes	No	Male

Yes	Yes	Yes	Yes	Female
Yes	Yes	No	Yes	Male
Yes	Yes	No	No	Female
Yes	Yes	No	Yes	Female
Yes	Yes	No	No	Male
Yes	Yes	Yes	Yes	Female
Yes	No	No	Yes	Female
Yes	Yes	Yes	Yes	Female
Yes	Yes	No	Yes	Female
Yes	Yes	No	No	Female
Yes	Yes	Yes	No	Female









If these data are relied upon, then it is imperative to accept that instances such as criminal defamation and offence of enmity might arise if such a large percentage of the population fail to verify the authenticity of an AI generated photograph or video, thereby raising questions of who shall be held liable.

The question of liability and personhood of AI has already been discussed in certain case, of which a prominent case is that of a claim against an AI driven vehicle that injured a motor-cyclist. The plaintiff brought in a case

of negligence against the manufacturer. However, the defendant accepted its responsibility and there were no further in-depth deliberations on the question of legal personality and responsibility of the self-driven vehicle.[1] Patent applications with AI as the inventor have also been filed in the USPTO, but have so far been rejected but in South Africa, DABUS, a "creativity-machine" was granted patent in 2021. The createor of DABUS, Stephen Tahler has been fighting the USPTO over recognition of an AI as an inventor. Therefore, there is a lack of consensus about the legal personality of an AI and the approach would differ even more when it comes to questions of criminal liability, and delivery of justice.

4. Conclusion

Moral liability in itself is a prejudiced concept as it is affected by various factors, and hence rather it would be prudent to develop a standardized global ethical code, as has been proposed by the UNESCO. If a natural person holds the onus of culpability for actions of the AI, then determining liability would be easier. However, this is a matter that would need deliberation over with a specialized committee that can provide the technical know-how of the workings of an AI. Moreover, another approach could be adopted wherein person-hood can be granted to an AI pertaining to certain subject matters, however, criminal culpability and personhood of AI should be kept separate from one another, as there could be failure of delivery of justice and the developer of an AI and a person making use of an AI could act irresponsibly through a human-like machine without consequences. While the research takes into consideration only two criminal offences, it is imperative to understand that the development of AI has a huge potential and along with its development the activities that it would be capable of would be limitless. Therefore, a check and balance system have to be put in place in the race between technology and effective legislation, so as not to lag behind in appropriate legislation. ChatGPT has reignited the fire of the debate of whether the age of AI would be a boon or a bane, and while discussions on this are legitimate, the basic principle of the criminal justice system is that *mens rea* and *actus reus* must be established so as to hold a "person" liable for the any criminal act. While the manufacturer accepted responsibility in the AI-driven car case, it stole from us the opportunity of a real-life debate on whether there was foreseeability on the part of AI of such an accident occurring, if so, could it make informed decisions that were guided by morality and ethics to execute a decision.

Refrences:

1. George, A. (2023). Thwarting Bias in AI Systems. Carnegie Mellon University. https://engineering.cmu.edu/news-events/news/2018/12/11-datta-proxies.html
2. Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and Information Technology, 6(3), 175-183.
3. Bhalla, A. v. Chawdhury, S. & Ors (CRL.M.A. 8466/2019).
4. Williams, C. (2023). Can Machines be Morally Responsible. Prindle Post. https://www.prindleinstitute.org/2022/05/can-machines-be-morally-responsible/
5. Barsky, D. J. (2023). Artificial Intelligence Key Legal Issues: A Practical Guidance Practice Note. Lexis Nexis Practical Guidance. https://www.hklaw.com/-/media/files/insights/publications/2023/06/artificial-intelligence-key-legal-issues.pdf
6. Santoni de Sio, F., & Mecacci, G. (2021). Four Responsibility Gaps with Artificial Intelligence: Why They Matter and How to Address Them. Philosophy & Technology, 34, 1057-1084. https://link.springer.com/article/10.1007/s13347-021-00450-x
7. Chaudhary, G. (2021). Artificial Intelligence: The Personhood Conundrum. Artificial Intelligence and Law. https://ssrn.com/abstract=3804265
8. Matthias, A. (n.d.). The responsibility gap: Ascribing responsibility for the actions of learning automata. Retrieved fromhttps://philpapers.org/rec/MATTRGm3#:~:text=The%20responsibility%20gap%3A%20Ascribing%20responsibility%20for%20the%20actions,legally%20responsible%20for%20the%20consequences%20of%20its%20operation.
9. Ministry of Information. (2021). Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("IT Rules 2021").
10. Gordon, J. S. (2021). AI and Law: Ethical, Legal, and Socio-Political Implications. AI & Society, 36, 403-404. https://link.springer.com/article/10.1007/s00146-021-01194-0#citeas
11. Tarone, K. v. State of Maharashtra and Anr (APL-573.16 (1).odt-Judgment).
12. Lewis, T. (2014). A Brief History of Artificial Intelligence. Live Science. https://www.livescience.com/49007-history-of-artificial-intelligence.html
13. Nilsson v. General Motors LLC, Case No. 4:18-cv-00471.
14. Omowple, A. (2022). Research Shows AI Is Often Biased. Here's How to Make Algorithms Work for All of Us. World Economic Forum. https://www.weforum.org/agenda/2021/07/ai-machine-learning-bias-discrimination/
15. Omowple, A. (2022). Research Shows AI Is Often Biased. Here's How to Make Algorithms Work for All of Us. World Economic Forum. https://www.weforum.org/agenda/2021/07/ai-machine-learning-bias-discrimination/
16. Salomon v A Salomon & Co Ltd [1896] UKHL 1.
17. United Nations Educational, Scientific and Cultural Organisation. (2022). Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO.