International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

The Legal Dilemma of Deepfakes Ai Liability and the Challenges of Digital Identity Theft

Ganesh Subramanian[1], Swathi S[2]

[1,2]Student, B.Com.LLB (Hons)

ABSTRACT

The European Union's Artificial Intelligence Act (EU AI Act) defines a "Deep Fake" as an "AI-generated or manipulated image, audio, or video content that mimics existing persons, objects, places, or other entities or events, creating a false appearance of authenticity or truthfulness to viewers." In other words, Using machine learning algorithms, the bogus videos are transformed into realistic videos. Several AI-related catastrophes, including automated car accidents that resulted in physical injuries, robots causing harm to labourers, AI-assisted online privacy assaults, AI-assisted fraud that involves face, speech, or signature imitations, AI-assisted digital fingerprints falsely classifying innocent people as criminals at airports, and AI-assisted fraud in elections, have made headlines. Cryptographic signing and the hashing of a video into a fingerprint are used to confirm and reconfirm whether a video came from its original source. Many strategies for detecting deepfake videos have been developed over the last decade as a result of the rising threat of these videos. However, the fundamental issue with such procedures is that they are inaccurate and time-consuming. It also brings additional concerns such as misinformation and disinformation. The deep fakes raise concerns regarding digital identity theft. Identity theft, a criminal offence under the laws of most countries, including India, is addressed in the digital context under the Information Technology Act. However, as deepfakes are generated using artificial intelligence, the issue of fixing liability becomes increasingly complex. Commentators have argued that existing legal liability concepts may fail to address future conflicts involving AI systems. The challenge with the current legal system is primarily posed by AI systems that function without human interaction, but also by AI systems that operate with minimal human assistance. The current legal framework does not clearly define whether artificial intelligence can be regarded as a legal person, which complicates the issue of determining the extent of AI liability for the creation and dissemination of deepfakes. This ambiguity presents significant risks to the protection of digital identities, highlighting the need for a thorough review of existing laws to ensure they are adequately equipped to tackle the unique challenges posed by AI-generated content.

Keywords: Artificial Intelligence, AI liability, Deepfakes, Legal Person, Identity Theft

INTRODUCTION

The shift from the pre-digital age to the technologically advanced and digital era is marked by numerous developments, one of the most significant being the rise of Artificial Intelligence (AI). By 2030, the AI market is projected to reach a value of $738.80 billion[1]. Artificial intelligence (AI) has made remarkable

[1] Anamika., The Legal Battle against Deep fakes: Copyright implications, DRM Mechanisms and Regulatory Perspectives in the Digital Age 1035 JCLJ (2024).

advancements across a variety of industries, including healthcare, education, the automotive sector, and most notably, the service industry. One significant innovation emerging from AI is the development of deepfakes, which utilise deep learning algorithms to produce highly realistic synthetic content. Deepfakes are generated using a technique called Generative Adversarial Networks (GANs), which involve two distinct neural networks: the generator, responsible for creating synthetic content, and the discriminator, which evaluates the authenticity of the generated content. The generator's role is to fabricate altered or synthetic content, while the discriminator evaluates and attempts to distinguish between genuine content and that produced by the generator. Through this adversarial process, the generator gradually improves, producing deepfakes that are nearly indistinguishable from real media over time.

While deepfakes present significant risks when misused, they also hold the potential to offer groundbreaking applications across various fields if utilised responsibly. For example, AI technology has been employed to digitally recreate the likeness of James Dean, a film actor who passed away in 1955, for an upcoming film titled "Back to Eden." This demonstrates how deepfake technology can be leveraged in the entertainment industry to bring historical figures back to life on screen. With the rapid evolution of AI and machine learning, the process of training models for generating deepfakes has been significantly simplified.[2] Today, several companies even offer services that enable users to upload digital data of deceased loved ones, creating AI-driven "deadbots" capable of engaging in simulated conversations with the living.[3] However, this advancement brings with it significant disadvantages when exploited improperly, particularly in eroding public trust in AI. This scepticism intensifies in instances where AI technologies contribute to criminal activities, especially in the absence of adequate redress mechanisms to compensate victims of AI-related crimes, unlike traditional crime compensation frameworks.

BACKGROUND OF THE STUDY:

The origins of deepfake technology can be traced to early efforts in the 2010s to generate realistic human images using computer-generated imagery (CGI). This progress eventually led to the development of Generative Adversarial Networks (GANs), a key machine learning concept. The term "deepfake" was coined in 2017 by a Reddit moderator, originally referring to pornographic videos created using face-swapping technology.[4] We may have encountered Instagram reels featuring audio manipulations where songs originally performed by other singers are altered to appear as though they are sung by public figures, such as Prime Minister Narendra Modi or various actors. This form of audio manipulation constitutes a type of deepfake technology. Certain manipulated videos or audio files may be easily detectable, while others may appear highly realistic. A survey conducted by cybersecurity company McAfee revealed that approximately 75% of Indians encountered content generated using deepfakes in the past 12 months. Deepfakes are generated through deep learning techniques, particularly Generative Adversarial Networks (GANs), which is the basis for the term "deepfake." This technology alters the original source audio, video, or images by swapping them with the target i.e., another individual. While deepfakes have the potential to disseminate misinformation and cause considerable harm, they can also be utilized for legitimate purposes

[2]Konstantinos Liakopoulos, Exploring the Potential Benefits of Deepfake Technology, MEDIUM (Aug. 13, 2024), Exploring the Potential Benefits of Deepfake Technology | by Konstantinos Liakopoulos | Medium.

[3]S.J. Velasquez, How AI Is Bringing Film Stars Back from the Dead, BBC (July 19, 2023), https://www.bbc.com/future/article/20230718-how-ai-is-bringing-film-stars-back-from-the-dead.

[4]Gabe Regan, A Brief History of Deepfakes, REALITY DEFENDER (June 1, 2024), https://www.realitydefender.com/blog/history-of-deepfakes.

that do not lead to negative outcomes. Deepfake technology holds the potential to revolutionise immersive learning experiences in the future, transforming traditional educational methods from textbooks to interactive videos—imagine, for instance, Isaac Newton himself explaining the law of universal gravitation. However, without appropriate regulations, the risks associated with this technology become significant. The World Economic Forum has recognized disinformation as one of the primary risks for 2024, with deepfakes being highlighted as one of the most troubling applications of AI.[5] If left unchecked, deepfakes could serve as powerful tools for manipulation and exploitation, leading to potentially harmful consequences.

LITERATURE REVIEW

The global impact and potential dangers of deepfake technology have been extensively studied, with many researchers highlighting the lack of consistent regulation and the question of AI liability when deepfakes are used to commit crimes.

In the paper titled "Deepfakes and Copyright Law: Inadequacy of Present Laws in Addressing the Real Issues," the authors employed a qualitative research methodology to investigate the legal implications of deepfakes. The study identifies deepfakes as a source of
various threats, including national security risks, political manipulation, and pornographic exploitation, the latter being the most prevalent. The paper conducts an in-depth analysis of existing legal frameworks and highlights the interplay between intellectual property law, tort law, criminal law, and civil law when dealing with deepfakes. Several legal remedies available to individuals affected by deepfake content include the right to privacy, the tort of breach of confidence, passing off, defamation, malicious falsehood, copyright infringement, performers' rights, data protection, and both civil and criminal remedies. However, the study notes that no specific criminal statute or civil liability regime currently exists to outlaw the creation or distribution of deepfakes explicitly. Additionally, when deepfakes involve deceased individuals, legal remedies such as the right to privacy, defamation, and malicious falsehood are unavailable. The paper emphasises the need for a comprehensive legal framework to protect individual privacy from the misuse of deepfake technology.[6]

“Artificial intelligence and legal liability: towards an international approach of proportional liability based on risk sharing” is a research paper by Mohammad Bashayreh and others which demonstrates that the existing liability principles may create a void when the problem at hand involves autonomous machines. This paper also advocates that a new international body should be enforced to ensure uniform application of principles and risk assessments in Artificial Intelligence (AI) applications[7].

“Education Regarding Impact of AI on Cybercrimes and Liability for AI” by Dr. Anusuya Yadav is a Conceptual Legal Research paper that elaborates on the role of Artificial Intelligence in cybercrimes and the relevant cyber laws regulating such acts and also the limitations of deploying on a wide scale[8].

---

[5] Anna Maria Collard, 4 Ways to Future-Proof Against Deepfakes in 2024 and Beyond, WORLD ECONOMIC FORUM (Feb. 12, 2024), 4 ways to future-proof against deepfakes in 2024 and beyond | World Economic Forum.

[6] Aranya Nath & Sreelakshmi B., Deepfakes on Copyright Law: Inadequacy of Present Laws in Determining the Real Issues, 15 INDIAN J. L. & JUST. 1 (2024).

[7] Bashayreh, M., Sibai, F. N., & Tabbara, A. (2020). *Artificial intelligence and legal liability: towards an international approach of proportional liability based on risk sharing. Information & Communications Technology Law*, 30(2), 169–192. https://doi.org/10.1080/13600834.2020.1856025

[8] Dr. Anusuya Yadav. *Education Regarding Impact of AI on Cybercrimes and Liability for AI, 58(5) PSYCHOLOGY AND EDUCATION* (2021)

"Criminal Responsibility of Artificial Intelligence Committing Deep Fake Crimes in Indonesia" by Asri Gresmelian Eurike Hailtik and Wiwik Afifah used the research methods Statute approach, conceptual approach, and comparative approach to determine the criminal liability of the offenders creating deepfakes using AI in Indonesia[9].

In the article "Have We No Decency? Section 230 and the Liability of Social Media Companies for deepfake videos" by Nicholas O'Donnell analysis involves a doctrinal approach, the author examines the intersection of deepfake technology and the legal responsibilities of social media platforms. The study highlights how deepfakes are generated by leveraging a large database of publicly available images from sources such as Google, stock photo repositories, and YouTube videos. Advanced deep learning algorithms are employed to manipulate these images by integrating facial data from celebrities into pornographic content. After training, the algorithm autonomously completes the manipulation with minimal human involvement. The paper recommends that regulatory bodies introduce new regulations or amend existing laws to hold social media platforms accountable for the distribution of deepfakes. The underlying assumption is that the prospect of liability would compel these platforms to actively prevent the unchecked spread of such videos by developing and deploying more advanced detection and removal technologies. However, a significant gap in this research lies in its failure to address the potential liability of the AI itself, especially considering the software's ability to generate deepfakes without further human intervention after the training phase.[10]

RESEARCH PROBLEM

The legal ambiguity concerning the status of artificial intelligence complicates the attribution of liability and the mechanisms for securing compensation for victims of harmful deepfakes.

RESEARCH OBJECTIVE

1. The paper critically examines the legal status of Artificial Intelligence, exploring the extent to which AI can be recognized as a legal person.
2. It analyzes the attribution of liability for AI's generation of derogatory or harmful deepfake content.
3. The paper evaluates the existing legal framework governing the prevention of harmful deepfake content.
4. It examines the regulatory mechanisms in place to control the creation and dissemination of deepfake content.

RESEARCH HYPOTHESIS

The lack of clear legal boundaries for Artificial Intelligence as a legal entity will make it difficult to assign blame for harm produced by deepfakes, hindering legal redress for the victims of identity theft.

RESEARCH QUESTION

1. Is Artificial Intelligence (AI) eligible to be considered as an artificial person?

[9] Asri Gresmelian Eurike Hailtik and Wiwik Afifah., *Criminal Responsibility of Artificial Intelligence Committing Deep fake crimes in Indonesia*., 2 Asian Journal of Social and Humanities (2024)

[10] Nicholas O'Donnell, *Have We No Decency? Section 230 and the Liability of Social Media Companies for Deepfake Videos*, 2021 U. Ill. L. Rev. 701.

2. How can the liability be shifted from the person using it to generate such content to Artificial Intelligence (AI)?
3. What are the existing protections available to the victims of such infringements?

AI LIABILITY

As artificial intelligence (AI) technologies become more widely used in industries such as healthcare, finance, and transportation, concerns about their liability have emerged. The emergence of AI has transformed how we approach complicated problems, providing an unprecedented degree of efficiency and precision. However, this technical breakthrough creates new hurdles, particularly in determining accountability when AI systems inflict harm or make errors. Traditional legal frameworks fail to keep up with the rapid breakthroughs of AI, demanding a reevaluation of existing liability rules.

ARTIFICIAL PERSON

The question of "Can a Machine Think ?" dates back to 1950 where A. M. Turning published a paper where he created a framework for discussing Artificial Intelligence[11]. Artificial Intelligence Systems (AIS) differ from conventional computer algorithms in that they possess the ability to learn autonomously, acquire knowledge, and develop solutions based on the evaluation of various factors, independent of the choices made by the developer or programmer[12].

Non-human beings shall be regarded as legal persons under the law. This is a distinct and fictional creation of the law called legal personality[13]. The United States Supreme Court has determined that a corporation is a legal person with no physical form, living solely in the 'Contemplation of Law'[14]. According to F. H. Lawson, all that is required for the existence of a [legal] person is for the legislature, judge, jurist, or even the general public, to decide to consider it as a matter of rights or other legal relations. In India, idols, rivers, corporations, rivers, animals, and texts are separate legal entities with their own rights and responsibilities. While they cannot be held directly liable, the doctrines of trusteeship and vicarious liability will be used to enforce their obligations. The idea of trusteeship or vicarious liability, however, cannot be applied to artificial intelligence because AIs are known to deviate from the goals of their creators or inventors, emphasising this in his paper, Laurence B Solum argues that AI should be treated as a separate legal entity[15]. As a result, the inventor cannot be held responsible for the activities of artificial intelligence. We will go into more detail about liability in a minute, but first, we need to understand the concept of Corpus and Animus in relation to artificial intelligence on how AI may control and hold assets (corpus) and have the intention or capability to manage such assets (animus).

Artificial intelligence holding property poses a severe danger to investments since it can forecast future market trends, which can be unfair and ethically wrong in many circumstances. Additionally, the developer of AI may hide behind the veil of artificial intelligence.

---

[11] A. M. TURING, I.—COMPUTING MACHINERY AND INTELLIGENCE, Mind, Volume LIX, Issue 236, October 1950, Pages 433–460, https://doi.org/10.1093/mind/LIX.236.433

[12] Paulius Čerka, Jurgita Grigienė & Gintarė Sirbikytė, Is It Possible to Grant Legal Personality to Artificial Intelligence Software Systems?, 33 COMPUTER LAW & SECURITY REVIEW 685 (2017), https://doi.org/10.1016/j.clsr.2017.03.022.

[13] Legal Personality of Non-Human Beings, TOPPR (n.d.), https://www.toppr.com/guides/legal-aptitude/jurisprudence/legal-personality-of-non-human-beings/.

[14] Trustees of Dartmouth College v. Woodward, 17 U.S. 518 (1819)

[15] Lawrence B. Solum, Legal Personhood for Artificial Intelligences, 70 N.C. L. REV. 1231 (1992), Illinois Public Law Research Paper No. 09-13, https://ssrn.com/abstract=1108671.

The European Union (EU) has accepted artificial intelligence as an artificial legal person and has established a precedent for world jurisprudence by becoming the first government to regulate artificial intelligence. The AI Act establishes a standardized framework across the European Union, grounded in a forward-looking definition of AI and a risk-based approach.

1. Minimal risk: Most AI technologies, such as spam filters and AI-powered video games, are not subject to mandatory requirements under the AI Act. However, businesses may choose to implement additional conduct guidelines voluntarily.
2. Specific transparency risk: Systems like chatbots are required to inform consumers that they are interacting with a machine, and certain AI-generated content must be clearly labeled as such.
3. High risk: High-risk AI systems, such as AI-driven healthcare applications and those used in recruitment, must comply with stringent requirements. These include risk mitigation measures, the use of high-quality data sets, clear user data protocols, and human oversight, among other obligations.
4. Unacceptable risk: For example, artificial intelligence (AI) systems that allow governments or enterprises to conduct "social scoring" are seen as an imminent danger to people's fundamental rights and are thus prohibited[16].

If AI is recognized as a legal person, it would be subject to liability for its actions in a manner similar to that of corporations or organisations. The EU impact assessment report advocates for a strict liability framework regarding claims arising from AI-related incidents, which also mandates the requirement for insurance coverage ensuring that victims receive adequate compensation for any damages incurred. The strict liability principle eases the burden of proof for victims, allowing them to establish instances of non-compliance without needing to demonstrate fault or intent.[17] The application of strict liability is contingent upon the risk classification of the AI system; for systems designated as high-risk, courts may presume non-compliance, shifting the burden of proof from the victim to the perpetrator. The 2019 Report outlines two categories of potential perpetrators: the "frontend operator," a natural or legal person who controls the risks associated with the AI system's operation and benefits from its use, and the "backend operator," a natural or legal person who defines the technological specifications, supplies data, and provides essential backend support services, thus controlling the risks of the AI system. Liability for both operators is assessed on a case-by-case basis, based on the degree of control they exert. The European Parliament Resolution of 2020 establishes a strict liability regime for operators, obligating them to compensate for damages when the actual perpetrator cannot be identified.[18]

While India has not directly adopted any specific laws from the European Union, several Indian laws have been influenced by EU legislative frameworks. Notably, various provisions of the Competition Act, of 2002, are inspired by elements of the EU Competition Law, reflecting the influence of European regulatory principles on India's legal system. The framework for regulating anti-competitive actions in India is based on concepts similar to those found in EU competition law, with an emphasis on fair competition, market

[16] https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_en?form=MG0AV3

[17] https://commission.europa.eu/document/download/a25ea208-9a1d-483b-ab71-bcd1905e9000_en?filename=1_4_197608_impact_asse_dir_ai_en.pdf#page=44.21

[18] A.T. da Fonseca, E. Vaz de Sequeira & L. Barreto Xavier, Liability for AI Driven Systems, in H. Sousa Antunes, P.M. Freitas, A.L. Oliveira, C. Martins Pereira, E. Vaz de Sequeira & L. Barreto Xavier (eds), Multidisciplinary Perspectives on Artificial Intelligence and the Law, Law, Governance and Technology Series, vol. 58 (Springer, Cham 2024), https://doi.org/10.1007/978-3-031-41264-6_16.

dominance, and anti-cartel measures[19]. Several EU laws and directives have influenced India's legal structure, particularly in the areas of data privacy, intellectual property, consumer rights, environmental protection, and competition law.

AI-GENERATED AND AI-ASSISTED

If a work embodying traditional aspects of authorship is created by an automated system, it lacks the element of human authorship and, therefore, cannot be registered with the Copyright Office. When AI technology gets a cue from a human and creates complicated literary, visual, or musical creations, it determines and executes the "traditional elements of authorship" rather than the human user.

To determine the liability of AI, it is essential to evaluate the two tests designed to assess the creativity of AI systems: the Turing Test and the Lovelace Test. The Turing Test, introduced by Alan Turing in 1950, is a thought experiment that suggests if an individual cannot distinguish between a human and an automated system through text-based interaction, the system must be considered intelligent. However, Turing did not intend for the test to be physically conducted; rather, he aimed to demonstrate that computers could exhibit human-like characteristics, despite the challenge of defining those traits.

In 2001, the Lovelace Test was proposed as an update to the Turing Test. It emphasizes that to assess human-like qualities in AI, one must consider the role of creativity, as human creation requires intelligence. The Lovelace Test proposes that an AI be tasked with generating something, such as a narrative or a poem, and that the test is passed only if the AI's programmer is unable to explain how the AI arrived at its output, thereby positioning creativity as a proxy for intellectual ability.[20]

According to the United States Copyright Office, understanding of the present generative AI technologies, users have limited creative influence over how such systems perceive prompts and develop work[21]. In Thaler v. Perlmutter, the court ruled that there were no ownership rights for creations generated by an AI tool developed by the plaintiff, Stephen Thaler. This decision was made despite Thaler's intentional limitation of human involvement in the creative process and his emphasis on the machine's role in generating the work.[22]

Although the recent AI Act in Europe does not directly address the ownership of AI-generated content, the Court of Justice of the European Union has offered some directional guidance in Infopaq International A/S v Danske Dagblades Forening. In this case, the court ruled that copyright protection applies only if the work demonstrates originality stemming from the "author's own intellectual creation," a concept that is often interpreted to require significant human involvement. In a similar vein, the United Kingdom's Copyright, Designs and Patents Act 1988 (CDPA) extends copyright protection to "computer-generated works," acknowledging that such works may be eligible for copyright, even if no human author is involved in their creation.

[19]Millia Dasgupta, How Different Is European Competition Law from Competition Law in India?, iPleaders (Feb. 20, 2021), https://blog.ipleaders.in/different-european-competition-law-competition-law-india/?form=MG0AV3.

[20]Sean O'Neill, How Creative Is Your Computer? The Lovelace Test Is a Better Measure of Artificial Intelligence Than the Turing Test, SLATE (Dec. 21, 2014), https://slate.com/technology/2014/12/lovelace-test-of-artificial-intelligence-creativity-better-than-the-turing-test-of-intelligence.html.

[21]Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 FR 16190 (Mar. 16, 2023) (to be codified at 37 CFR 202).

[22] Thaler v. Perlmutter, Civil Action No. 22-1564 (BAH) (D.D.C. filed Nov. 2022).

INDIAN SCENARIO

Deepfakes have become more prevalent in India, particularly during election cycles, where they are used to spread false or misleading information about political parties or their members, leading to confusion and eroding public trust. The World Economic Forum has recognized misinformation as one of the most significant threats during election periods.[23] To address this, the Misinformation Combat Alliance set up a Deepfakes Analysis Unit in March 2024, allowing people to report suspicious AI-generated content via WhatsApp. This initiative has already uncovered numerous harmful and misleading audio and video deep fakes including a manipulated audio falsely attributed to Tamil Nadu Chief Minister M.K. Stalin.[24]

India currently lacks dedicated laws specifically targeting deepfakes. Consequently, it is crucial to analyse the existing legal frameworks that, although not directly focused on deepfakes, can be applied to regulate their production and distribution. n the case of Rajat Sharma v. Union of India, a Public Interest Litigation (PIL) was filed by Mr. Rajat Sharma, a prominent journalist, before the Delhi High Court. The petition addressed a deepfake video in which he was falsely depicted as offering medical advice, raising concerns about the potential harm caused by such misleading content. Mr Sharma contended that deepfakes, particularly those involving public figures, pose a serious threat to society due to the potential for misinformation and manipulation. He argued that, as a prominent figure frequently seen on television, such deepfakes could have a heightened impact, as the general public is more likely to believe and rely on the false information disseminated in the video.

Mr. Sharma further submitted that deepfakes infringe upon the fundamental rights of individuals, including the right to privacy and reputation. He urged the court to direct immediate governmental action to address the growing menace of deepfakes by establishing a comprehensive legal framework to regulate their creation and dissemination. In response to the petition, the Delhi High Court directed the Union government, specifically the Ministry of Electronics and Information Technology (MeitY), to provide a response regarding the development of regulatory mechanisms aimed at addressing the issue of deepfakes.[25]

In the case of Lawyer's Voice v. Union of India & Ors., a public interest litigation (PIL) was filed concerning a deepfake video circulated on social media that portrayed India's Home Minister making controversial statements during the election period. The Delhi High Court noted the rising prevalence of deepfake videos targeting public figures, which can severely damage reputations and incite public outrage. The plea specifically sought the removal of such content from social media platforms, arguing that it undermined free and fair elections. However, the court dismissed the PIL, directing the Election Commission of India (ECI) to take appropriate action on the matter.[26] In response, the Election Commission of India (ECI) set up a "Myth vs Reality" registry, a verification mechanism designed to ensure the accuracy and credibility of information circulated during elections.[27]

[23]Global Risks Report 2024, WORLD ECONOMIC FORUM (Jan. 10, 2024), https://www.weforum.org/publications/global-risks-report-2024/.

[24]Stalin's Voice Clone or Impression Used for Fake Narrative About Karunanidhi, DEEPFAKES ANALYSIS UNIT (Aug. 5, 2024), https://www.dau.mcaindia.in/blog/stalins-voice-clone-or-impression-used-for-fake-narrative-about-karunanidhi.

[25] Rajat Sharma v. Union of India Writ Petition(s)(Civil) No(s). 80/2021

[26] Lawyer's Voice v. Union of India & Ors 2024 LiveLaw (Del) 534

[27] ELECTION COMMISSION OF INDIA, https://mythvsreality.eci.gov.in/.

DEEPFAKE AND INTELLECTUAL PROPERTY RIGHTS

Intellectual property laws, which are primarily designed to safeguard original creative works, do not explicitly address deepfakes. However, the infringement provisions within these laws can be interpreted to encompass and restrict the distribution of deepfakes, particularly where such content involves unauthorised use or manipulation of protected works.

Deepfakes often involve the use of copyrighted material, leading to potential infringements on the rights of copyright holders. This raises the question of whether the Indian Copyright Act is sufficient to address the challenges posed by deepfakes and other forms of AI-generated content, particularly in terms of protecting individuals' rights and preventing the spread of misinformation. Section 51 of the Act comprehensively outlines the circumstances that constitute copyright infringement, specifically when the act is done without the authorisation of the owner[28]. A deepfake involves the manipulation of photos or videos, and if created by utilising someone else's property without permission, such unauthorised use constitutes an infringement upon the rights of the original owner.[29]

Conversely, Section 52 of the Indian Copyright Act outlines exceptions to copyright infringement, one of which is the doctrine of "fair dealing."[30] The fair dealing exception under Indian law could be relevant in cases where deepfakes are created with malicious intent. However, the Indian doctrine of fair dealing is comparatively rigid when contrasted with the broader "fair use" doctrine in the United States, which allows for reproduction in several instances, particularly when serving the public interest. Given this distinction, while Indian Copyright Law can potentially be applied to the creation and distribution of deepfakes, its narrow approach to exceptions, such as fair dealing, may limit its efficacy in addressing the complexities posed by deepfake technology. However, the rigidity of the Indian Copyright Act, which restricts fair dealing to statutorily defined categories, may not sufficiently protect the legitimate and fair use of deepfakes for purposes such as entertainment.[31] The limited scope of fair dealing under Indian law fails to account for scenarios where deepfakes might be used in a legitimate, non-malicious manner, particularly in creative industries. Consequently, additional legal frameworks may be necessary to regulate deepfake technology effectively.

PERFORMER'S RIGHTS

Section 2(qq) of the Indian Copyright Act defines a performer as encompassing "actors, singers, musicians, dancers, acrobats, jugglers, conjurers, snake charmers, lecturers, and any individuals who engage in a performance". This inclusive definition extends to lecturers, politicians, journalists, and public figures who share their performance recordings on social media platforms.[32] Consequently, using these recordings as input data to generate deepfakes would fall within the scope of performers' rights, and any unauthorized use of such material could potentially constitute a violation of those rights.

[28] Section 51 of The Copyright Act, 1957

[29] Shinu Vig, Regulating Deepfakes: An Indian Perspective, 17 J. STRATEGIC SECURITY 5 (2024), Regulating Deepfakes: An Indian perspective

[30] Section 52 of The Copyright Act, 1957

[31] Sindhu A, Interventions on the Issue of Deepfakes in Copyright, (2023), Interventions on the Issue of Deepfakes in Copyright

[32] Pavis, Mathilde. "Rebalancing Our Regulatory Response to Deepfakes with Performers' Rights." Convergence, (2021). Accessed October 10, 2024. https://doi.org/10.1177/13548565211033418

PERSONALITY RIGHTS

Although the Copyright Act does not explicitly recognize personality rights, courts may, in certain instances, invoke copyright law to safeguard these rights, particularly when the use of an individual's likeness or performance infringes upon their personal interests. In the case of Anil Kapoor v Simply Life India and Ors[33] The celebrity's face was altered using Artificial Intelligence, resulting in the dilution and tarnishment of their public image. Similar incidents have occurred with actors Rashmika Mandanna and Sara Tendulkar, where AI-generated manipulated images were used to violate their image rights. The use of AI to create falsified or misleading images of celebrities is a relatively recent but increasingly common method of infringing on their rights. In other instances, defendants have also been involved in morphing actresses' images, including those of Katrina Kaif, Sridevi, and Madhuri Dixit. Ruling in favour of the plaintiff, the court acknowledged the challenges posed by emerging technologies like Artificial Intelligence, particularly dark patterns and deep fakes, which have heightened the risk of manipulation. Deepfakes, in particular, have the capability to morph a celebrity's face for commercial gain or to damage their reputation. The court emphasized the necessity for legislative measures to mitigate the risks posed by such technologies.

IDENTITY THEFT

Although there is no specific legislation in India that addresses deepfakes, both civil and criminal remedies are available under existing legal frameworks. The Information Technology Act, 2000, contains relevant provisions that can be invoked in such cases, including Section 66E, which addresses privacy violations, and Section 66D, which pertains to offenses related to impersonation or cheating committed through computer systems. Additionally, Sections 67, 67A, and 67B of the Act allow for the prosecution of individuals engaged in the publication or transmission of deepfakes that are deemed obscene or sexually explicit. Moreover, the Act imposes liability on intermediaries, such as social media platforms, under Section 79 for hosting unlawful content; non-compliance with this obligation may result in the forfeiture of their safe harbour protection. The Union government issued an advisory to social media intermediaries, requiring them to remove any reported deepfake content within 36 hours of receiving a complaint. The advisory further emphasised the obligation of these platforms to exercise due care in accordance with the due diligence rules, particularly to prevent the dissemination of deepfake misinformation. This includes ensuring that content violating platform rules, regulations, or user agreements is swiftly addressed, aiming to mitigate the spread of harmful and misleading information.[34] The current cybercrime legislation in India is insufficient to adequately address the issue of deepfakes. As the Information Technology Act, of 2000, lacks specific provisions concerning artificial intelligence, machine learning, or deepfakes, regulating the use of these technologies presents significant challenges.

The provisions of the Bharatiya Nyay Sanhita (BNS) offer legal remedies for cybercrimes involving deepfakes. Section 336, addressing forgery, and Section 319, which pertains to cheating by impersonation, are applicable in cases where deepfakes are used to manipulate identity or commit fraud. Additionally, Section 79 of the Information Technology Act, which addresses acts intended to insult the modesty of a woman, can be invoked when deepfakes are used to create harmful or pornographic content that violates an individual's dignity. For cases involving defamation through deepfakes, Section 356, dealing with

[33] Anil Kapoor v Simply Life India and Ors CS (COMM) 652/2023

[34] Ministry of Electronics & IT, Advisory to Social Media Intermediaries to Identify Misinformation and Deepfakes, (Nov. 7, 2023), https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1975445

criminal defamation, provides legal recourse to protect the reputation of the affected individual. In a recent incident involving the dissemination of a deepfake video featuring the renowned celebrity Rashmika Mandanna, the Delhi Police acted promptly by invoking various legal provisions. They invoked Sections 465 and 469 of the Indian Penal Code, which relate to forgery and forgery committed with the intent to harm an individual's reputation, to address the reputational damage caused by the manipulated content. Furthermore, the Information Technology Act, 2000, was also invoked, specifically Sections 66C, which addresses identity theft, and 66E, which concerns the violation of privacy through the unauthorized capture, publication, or transmission of an individual's image.[35]

PRIVACY CONCERNS

The creation of deepfakes can result in a violation of an individual's privacy, which is safeguarded under Article 21 of the Indian Constitution, as well as the Digital Personal Data Protection Act (DPDPA) of 2023. Personal data is broadly defined under Section 2(t) of the Act as any data concerning an individual who can be identified by or is related to such data, including a person's photograph or video.[36] These personal identifiers can be misused by Generative Adversarial Networks (GANs) to produce deepfakes.

Under Section 8(5) of the Act, data fiduciaries are obligated to protect personal data. In this context, if a social media platform restricts the downloading of personal data, it could hinder AI systems from being trained on such data, thereby potentially reducing the generation of harmful deepfakes. Moreover, the fiduciary duty to protect personal data can be interpreted to extend to the removal of deepfake content that persists on such platforms.

However, the DPDPA does not fully address the challenges posed by generative AI-based media. Section 3(c) of the Act provides exemptions for the use of data for personal or domestic purposes, a provision that may prove inadequate in effectively combating the misuse of data in creating deepfakes. Content generated for personal use can be easily disseminated in today's digital landscape, spreading rapidly like wildfire. Furthermore, the same section exempts publicly available data from protection, which leaves room for the creation of deepfakes using publicly accessible images, such as those of celebrities or politicians, thus posing a significant challenge in regulating harmful deepfake content.

RESEARCH METHOD

This research uses a systematic approach with conceptual legal research majorly relying on secondary sources.

SCOPE AND LIMITATION

SCOPE OF THE PAPER

1. Investigating if AI can be treated as a legal entity. This entails reviewing existing laws and cases to determine how the legal system currently perceives AI.

[35]Arvind Ojha, Rashmika Mandanna Deepfake: How Cops Traced Accused, Techie from Andhra's Guntur, INDIA TODAY (Jan. 20, 2024), https://www.indiatoday.in/india/story/rashmika-mandanna-deepfake-video-accused-arrested-andra-engineer-wanted-to-boost-followers-2491386-2024-01-20.

[36]Sarvagya Chitranshi, The "Deepfake" Conundrum: Can the Digital Personal Data Protection Act, 2023 Deal with Misuse of Generative AI?, 2023 IJLT (Dec. 23, 2023), https://www.ijlt.in/post/the-deepfake-conundrum-can-the-digital-personal-data-protection-act-2023-deal-with-misuse-of-ge.

2. The study goes into the challenging question of who is responsible when deepfakes cause harm. We'll look at both the creators of these technologies and the people who utilise them.
3. Examining the current laws governing deepfakes and identity theft, determining how effective they are at preventing harm and addressing issues that arise.
4. Looking into what protections are available to victims of deepfakes, such as compensation and legal recourse.

LIMITATIONS

1. Because AI is advancing rapidly, new innovations may arise after our research is completed, thereby affecting the applicability of our findings.
2. This paper focuses solely on specific legal systems, such as those in India or the EU. This means that our findings may not apply generally in other nations with different laws.
3. We have not explored the technical aspects of AI and deepfake technologies, As this could lead to some oversimplifications of the legal ramifications.
4. The ethical difficulties surrounding deepfakes might be subjective, with differing perspectives among organisations. This may alter the way we interpret our findings.
5. Limited access to data on the effect of deepfake-related damages, limit our ability to form firm judgments.
6. While legal concerns were discussed in depth, this research did not comprehensively address broader societal aspects, such as public confidence in AI-generated content and the rapid dissemination of disinformation.

RECOMMENDATIONS

The lack of sufficient regulatory frameworks to address deepfake technology poses significant risks, with the potential for serious and far-reaching consequences. These can range from inciting public disorder to exploiting innocent individuals for financial gain. For example, in a 2023 incident in Kerala, a deepfake video was used to convincingly replicate the voice of the victim's colleague, fraudulently requesting financial assistance for a relative in need. Trusting the authenticity of the video, the victim transferred ₹40,000.[37]As artificial intelligence continues to evolve, such scams are becoming more widespread, blurring the line between reality and fabrication, thereby making it increasingly challenging for individuals to distinguish between authentic and manipulated content.

The risks extend beyond individual fraud and present a real threat to the judicial system. A growing concern is the potential for deepfake technology to tamper with evidence presented in court. A recent case in the UK saw the submission of a deepfake audio recording, which was intended to discredit the father in a child custody dispute. This alarming scenario highlights the pressing possibility of similar manipulations within the Indian judicial system, which could undermine the very foundations of justice.[38] Given the growing sophistication of AI-generated content and the rapid spread of deepfakes across digital platforms, prompt and comprehensive regulatory measures are crucial to address the associated risks. Such legal safeguards are necessary for the trust in judicial systems, public institutions, and individual rights to be

[37] Amit Chaturvedi, Kerala Man Loses ₹40,000 to AI-Based Deepfake Scam: Here's What It Is, NDTV (Jul. 18, 2023),https://www.ndtv.com/india-news/kerala-man-loses-rs-40-000-to-ai-based-deepfake-scam-heres-what-it-is-4249684.

[38] Deepfake Audio Evidence Used in UK Court to Discredit Father, CYFOR BLOG https://cyfor.co.uk/deepfake-audio-evidence-used-in-uk-court-to-discredit-father/

significantly eroded. Therefore, this research paper recommends the following measures to address these concerns effectively:

1. Recognizing AI as a distinct legal entity may alter liability and responsibility in digital interactions. By providing a defined legal standing for AI, we can effectively address concerns resulting from AI-generated content, such as deepfakes, ensuring that victims obtain justice while also encouraging responsible research and usage of AI technologies.
2. Establishing a no-fault liability regime, instead of relying solely on strict liability, based on the gravity of the offence, would significantly enhance victim protection in cases involving AI-generated harm, such as deepfakes. The burden of proof would be shifted away from the victim, ensuring that operators cannot evade liability by merely proving they exercised due care. This approach would hold AI developers and operators accountable for harm caused, irrespective of fault, and ensure that victims receive appropriate redress without facing the evidentiary hurdles common in traditional liability claims.
3. A dedicated compensation fund or Mandating an insurance for AI systems could serve as an effective solution for providing compensation in cases involving harm caused by AI-generated content, particularly in instances where it is difficult or impossible to trace the origins of the AI-generated content or the software used to create it. In such cases, victims should not be left without legal remedies or adequate redress. AI-related offences must be treated on par with similar non-AI crimes where liability is clearly established, ensuring that victims receive fair compensation.
4. Labeling AI-generated content could serve as an effective safeguard to mitigate the risks associated with the misuse of artificial intelligence, especially in the context of deepfakes, by providing clear identification and reducing potential harm. Intermediaries could be mandated to label all AI-generated content and implement a self-regulatory review mechanism to assess and approve such content before dissemination. This would help in preventing the circulation of harmful or misleading media. Additionally, AI systems themselves could be required to incorporate watermarking or other identifiable markers to clearly indicate that the content is AI-generated. Such measures would not only enhance transparency but also raise public awareness, allowing individuals to distinguish between authentic content and potential deepfakes.
5. An authority at the Union level should be established to specifically address complaints arising from the misuse of deepfakes and other AI-generated content. This centralized authority would be entrusted with the responsibility of managing, investigating, and resolving legal matters pertaining to the unlawful use of such technologies. To facilitate effective enforcement, the authority should be empowered with statutory powers to engage with intermediaries, mandating them to take swift and decisive action, including the prompt removal of harmful deepfake content from their platforms.

CONCLUSION

In conclusion, examining the legal consequences of artificial intelligence and deepfake technologies reveals a complex terrain that requires immediate attention. As AI evolves and integrates into numerous industries, its potential for abuse, particularly in the creation of deepfakes, poses serious threats to individual rights, public trust, and societal norms. Also, the moral consequences of deepfake technology are significant. The risk of misinformation and a loss of faith in broadcasting and communication systems is growing. As deepfakes get more advanced, they may be employed to control public opinion, influence elections, and destroy reputations, all of which might have serious ramifications for democracy and

societal cohesiveness. AI's lack of a distinct legal position complicates culpability attribution, making it difficult to hold the parties responsible liable when deepfakes do harm. This ambiguity not only impedes the pursuit of justice for victims, but it also raises more general concerns about the prospect of AI governance. Furthermore, the study emphasises the significance of aggressive legislation to control the creation and transmission of deepfakes. Effective legal frameworks should not only prohibit criminal use, but also allow for quick reactions to occurrences involving AI. Legislators can help develop public trust in digital technology and promote responsible innovation by defining clear liability norms and victim protection measures. In a nutshell while deepfake technology presents exciting opportunities for creativity and innovation, its perils must not be underestimated. Addressing the legal difficulties and ethical concerns around AI and deepfakes is critical to ensuring that the advantages of these technologies are used responsibly, eventually protecting individual rights and sustaining public trust in the digital age.