Digital Age: Navigating Legal Landscape vis-à-vis Addressing Deepfakes and Manipulated Media

Journal of Development Economics and Management Research Studies (JDMS)
A Peer Reviewed Open Access International Journal
ISSN: 2582 5119 (Online)
Crossref Prefix No: 10.53422
12 (26), 22 - 29, October – December, 2025
@Center for Development Economic Studies (CDES)
Reprints and permissions
https://www.cdes.org.in/
https://www.cdes.org.in/about-journal/

Digital Age: Navigating Legal Landscape vis-à-vis Addressing Deepfakes and Manipulated Media

Kushala Simha[1]

Abstract

Deepfake technology presents a complex intersection with legal challenges, touching on issues ranging from privacy and intellectual property to national security and defamation. Initially, these technologies emerged as a form of entertainment, allowing users to swap faces in videos for humorous or creative purposes. However, the darker side of deepfakes soon became apparent as malicious actors began exploiting the technology for nefarious purposes. Existing privacy laws are ill-equipped to handle the nuances of deepfake creation and distribution, necessitating a re-evaluation of legal frameworks. Content creators may find their work repurposed in ways that were previously unimaginable, prompting legal battles to establish the limits of artistic and intellectual ownership in the age of deepfakes. Legal systems must adapt to distinguish between malicious intent and unintentional misinformation, balancing the protection of reputation with the right to freedom of expression.

Beyond individual harm, deepfakes pose serious threats to national security. The rise of deepfakes underscores the delicate balance required between fostering technological innovation and implementing effective regulations. Striking this balance necessitates international collaboration, adaptive legal frameworks, and ongoing dialogue between legislators, technologists, and civil society. The aim of this research paper is to investigate the issues pertaining to Privacy, Intellectual Property, Security, Defamation and Misinformation. As deepfake technology continues to evolve, the legal landscape must evolve in tandem to safeguard individual rights, privacy, and societal stability.

Keywords: Deepfake, Privacy, Intellectual Property, Security, Defamation, Misinformation.

[1] 4th Year- BA LLB (Hons), School of Law, Mahindra University.

INTRODUCTION:

Deepfakes are films that intentionally mislead people by using artificial intelligence (AI), deep learning, and Photoshop to produce pictures of real events that propagate false information. The movies are produced by the interplay of two technologies: machine learning (ML) and generative adversarial networks (GANs). That is "how the late actor Paul Walker's resurrection for Fast & Furious 7 came to be. Manoj Tiwari, a politician running for the Indian legislative assembly in 2020, had his speech translated into the "Haryanvi" dialect. To give the deepfake video a realistic understanding or touch, the designer must "first train a neural network on several hours of real video footage of the subject". After that, a copy is superimposed using a combination of computer graphics and the trained network. The "marketplace of ideas" theory—which held that access to more knowledge was preferable to excessive content censorship—was once the cornerstone of American civilization. In "Abrams v. United States" (Abrams V, 1919), Justice Oliver Wendell Holmes stated in his well-known dissent that free exchange of ideas in a cutthroat market serves as the finest barometer of truth. Holmes's remarks are poignant and sentimental in the age of fake news. If the media is full of misinformation and fake news, can society still rely on the "marketplace of ideas" for accurate content? The federal government should allow organisations skilled in disinformation to provide platform rules to direct their self-regulation, while politicians rush to find solutions to the manipulated media issue that frequently feeds viral misinformation.

While this objectionable content was initially referred to as a "deepfake" by a Reddit user in 2017 to describe movies that were altered to include a celebrity's face overlaid on already-existing pornography, deepfakes now attack cybersecurity, companies, and politicians equally. Interest in the new media is growing even though Reddit removed the deepfake thread at the beginning of 2018, as development tools become accessible to non-experts through computer apps and service portals. Although deepfakes have received most of the attention, there is a bigger plan, including manipulated media that is spreading misinformation on the internet. Because the internet is so widely used, a falsehood shared online can stick around. Some have claimed that "the right kind of algorithmically selected" material reinforces our fundamental prejudices, which is why manipulated media is dangerous. This powerful mixture of prejudices and misinformation on social media can make something "go viral." Remember "Pizzagate," the partisan conspiracy theory that originated on Facebook and quickly gained traction on Twitter, claiming that Hillary Clinton, the 2016 Democratic nominee for president, operated a child sex business out of a restaurant? Disinformation may have serious consequences, as seen by the fact that a guy armed with a.38 pistol, an AR-15 semiautomatic rifle, and a folding knife broke into the restaurant in Washington, D.C., despite the rumour having no basis in truth. Notwithstanding the violence, individuals now find it more difficult to believe sources of information that once seemed to be quite reliable. This threat is concerning government and business players. "A day after a lewd video of actor Rashmika Mandanna appeared on several social media sites, on November 07, 2023", she released a statement questioning its veracity. Unknowingly, the actor's face was overlaid on a British Indian influencer's body. Because artificial intelligence techniques were utilised to edit the photos and videos, this is an example of a "Deepfake" video. The newly produced photos are a type of misinformation, and the context and one's perception of them will determine whether or not one finds them offensive. Celebrities are easily targeted, and their offensive videos are a highly marketable product.

1. Role of Administrative Bodies:

Administrative bodies have the authority to control artificial intelligence (AI) and other synthetic or altered media that equate to unfair or deceptive commercial activities and practices. However, this authority is probably limited to media that advertise "food, drugs, devices, services, or cosmetics." Even though the government could seem like a better fit, it doesn't currently seem like that organisation has the authority or the will to control what is shared on social media. Finally, it should be noted that although the Administrative Agencies have the authority to control speech related to campaigns, they do not control the veracity of such speech and are unlikely to do so because of the logistical, political, and constitutional issues that would arise from doing so. There are financial regulations pertaining to elections. By superimposing human faces on obscene films, deepfake technology allows for the creation of explicit content without requiring permission. By utilising technology to create the material, this goal is met. The unauthorised use of people's identities raises considerable privacy concerns due to the substantial possibility of harm. To grasp the legal environment surrounding digital falsifications, it is important to grasp the last five factors. First, it is frequently difficult and expensive to find the persons who manufacture detrimental falsifications, which makes attempts to hold them accountable for their actions hampered by attribution issues. Second, and somewhat related, offenders frequently reside outside of the nation and may not be subject to the judicial system of that nation. Third, bringing civil claims can be costly and dangerous, and those who have been harmed by false information may worry that going to court will draw even more unwelcome attention—a phenomenon referred to as the Streisand effect. Fourth, the veracity of a digital falsification may determine legal accountability; nonetheless, this is a matter that varies from case to case. In the event that a deepfake depicted a presidential candidate making racist remarks, for instance, she would probably need to demonstrate that people had a reasonable belief that she made the remarks in order to successfully file a defamation lawsuit. The deepfake would probably be regarded as satire or parody by the courts if it were truly unbelievable. It is safe to assume that credible falsifications are both more likely to be legally problematic, even if the relevance of believability will vary depending on the type of legal claim and the specifics of each case. Deepfakes are already causing deeply troubling harm to regular people, especially when actors use their likenesses to make nonconsensual pornography. This abuse mainly targets women; victims have included minors, and it may make it easier for evil actors to extort and abuse more people. With the development of deepfakes, the non-consensual sharing of intimate imagery—also referred to as revenge porn—is growing, although it is ultimately not a new issue.

However, not all of the current laws sufficiently penalise malicious actors for disseminating or threatening to disseminate this material, nor do they always address content created by AI. Legislators ought to establish severe criminal and civil penalties for both distributing non-consensual intimate audio-visual content—including content produced by artificial intelligence—and for making threats to do so. In cases where the victim is a minor, the penalty ought to be extra harsh. By endorsing and passing the bipartisan Preventing Deepfakes of Intimate Images Act in the United States, legislators can take immediate action on this recommendation. This measure would allow affected parties to seek damages and impose liability on anyone who discloses or threatens to disclose an in consensual, intimate digital image of someone, including content generated by artificial intelligence. By establishing a much-needed federal baseline of accountability—which is unevenly addressed by state-level revenge porn laws—this proposal would give victims and people in the US more safety. Many of these issues are already addressed by the EU AI Act, which IBM has long supported. It covers deepfakes more broadly and imposes transparency rules that make it clear when specific information is not authentic. Policymakers should pay close attention to making sure people are shielded from non-consensual intimate audio-visual content as they anticipate the Act's implementation in the upcoming months.

2. IT Rules and India's Legal Regime:

India does not yet have a special law that addresses offences with deepfakes and artificial intelligence. Celebrities and other well-known individuals have occasionally used their personality rights—that is, the right to privacy and the right to publicity, respectively—to prevent the improper use of their likeness, voice, persona, and other characteristics in relation to deepfake content. More generally, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, as modified ("IT Rules"), and some sections under them may be helpful in this regard. The Information Technology Act, 2000 ("IT Act") and associated rules may also be of assistance. Nonetheless, there has been a legal challenge to the legality of these recently modified regulations. The intermediaries "lose their safe harbour protection under section 79 of the IT Act and shall be liable for consequential action or prosecution as provided under any law for the time being in force, including the IT Act and the Indian Penal Code, including section 469 of the IPC" if they fail to comply with the legal obligations outlined in the IT Rules, 2021. Deepfake films of celebrities like "Rashmika Mandanna, Nora Fatehi, Katrina Kaif, Kajol, and cricket player Sachin Tendulkar have been making the rounds on the internet in the past". "The IT Rules, 2021 cast specific legal obligations on intermediaries, including social media intermediaries and platforms, to ensure their accountability towards safe & trusted Internet including their expeditious action towards removal of the prohibited misinformation, patently false information and deepfakes," the minister of state for electronics and information technology.

"Intermediaries must comply with the orders of the Grievance Appellate Committee within the timeline mentioned in the order and publish a report," Chandrasekhar says. "Intermediaries must enable users, victims, or any person on their behalf, to also report violations relating to Rule 3(1)(b) or Rule 3(2)(b) (The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021) in a simple and easily accessible manner, including through in-app user reporting." It is recommended that intermediaries use further procedures to prevent the promotion of illicit loan and betting applications. Organisations must also set rules for the production and distribution of deepfake content by their staff members. Policies of this kind must also promote the organisation's internal and external deepfake technology usage in a responsible manner. It also reaffirmed that organizations risk losing the protection provided by Section 79(1) of the IT Act if they do not comply with the pertinent provisions of the Information Technology Act, 2000 (hereinafter referred to as the "IT Act") and Rules, Rule 7 of the Information Technology Rules (Intermediary Guidelines and Digital Media Ethics) Code, 2021 (hereinafter referred to as the "IT Rules"). Online intermediaries are not liable for any third-party data, information, or communication connection that they host or make available under Section 79 (1) of the IT Act. The Indian Penal Code's provisions may be invoked in court by anyone who feels wronged, according to Rule 7 of the IT Rules. A person who violates the privacy of another by posting or transmitting an image of their private area without that person's agreement faces up to three years in prison and a fine of INR two lakh, according to Section 66E of the IT Act. The IT Act's Sections 67, 67A, and 67 B expressly forbid and specify penalties for publishing or distributing pornographic content, content that features sexually explicit acts, and content that features children in such acts in electronic format. Social media companies have been encouraged to take action within 24 hours of receiving a complaint about any content in which there is electronic impersonation, including the use of electronically modified photos of individuals. Given this, Section 66D of the IT Act punishes "anyone who uses a communication device or computer resource" to deceive someone by impersonating them with a maximum fine of one lakh rupees and a term of three years in prison.

3. Comparative Study:

Deepfake legality is complicated in the US. Defamation claims can be made by victims; however, material removals could be interpreted as censorship and could therefore be against the First Amendment, which safeguards the rights to petition, assembly, expression, and religion. On the other hand, users have the ability to ask organisations like Google and Facebook that have gathered their data to remove it through the Right to be forgotten. Deepfakes have been used maliciously for things like face recognition system hacking and revenge porn. They erode public confidence in the media and muddy the distinctions between reality and fiction. Deepfakes can spread false information that people mistake for fact, which could cause societal upheaval. The use of deepfake technology is neither prohibited nor regulated by any particular laws or regulations in India. A worldwide framework for the development of "ethical" AI tools has been demanded by India. Certain features of deep fakes, like defamation and releasing sexual material, may be subject to existing regulations, such as Sections 67 and 67A of the Information Technology Act, 2000. Defamation is punishable under the Indian Penal Code (1860) Section 500. There is some protection against the exploitation of personal data thanks to the Digital Personal Data Protection Act of 2023. According to the Information Technology Rules, 2021, content that impersonates other people and photographs that have been electronically manipulated must be removed within 36 hours. India should create a thorough legal framework that targets deepfakes specifically, taking into account the possible effects on social stability, privacy, national security, and democracy. The first-ever AI Safety Summit 2023, which brought together 28 major nations, including the US, China, and India, decided that international action was required to address the possible threats associated with AI. The summit's "Bletchley Park Declaration" recognised the dangers of wilful misuse and losing control over artificial intelligence (AI) technologies.

December 2023 saw the holding of the Global Partnership on Artificial Intelligence (GPAI) meeting in New Delhi. The artificial intelligence-focused New Delhi Declaration was adopted as the conference came to an end. The proclamation-built consensus, GPAI members on creating safe, secure, and trustworthy AI and commitment to supporting the sustainability of GPAI initiatives. Tech businesses that sign up to the European Union's Code of Practice on Disinformation are required to combat deepfakes and fake accounts within six months of doing so. Tech corporations risk fines of up to 6% of their global annual revenue if found to be in violation. The first comprehensive legislation governing the use of artificial intelligence was passed by the European Union. A standard legislative and regulatory framework for artificial intelligence is intended to be introduced under the Artificial Intelligence Act (AI Act). The purpose of the draft rule is to guarantee the safety, respect for fundamental rights, and application of EU values of AI systems that are utilised in the EU and sold in Europe. To support the Department of Homeland Security in combating deepfake technology, the United States introduced the bipartisan Deepfake Task Force Act. China enacted extensive deep synthesis regulations that will take effect in 2023. The legislation mandates traceability of deep synthesis content and explicit labelling in an effort to combat misinformation. Under the Regulations, both suppliers and consumers of so-called "deep synthesis technology" are subject to requirements. Prominent digital corporations such as Meta and Google have declared plans to tackle the problem of deeply phoney material.

Nonetheless, their systems still have flaws that let this kind of stuff spread. Google has released watermarking and metadata as techniques for detecting fake content. While metadata gives original files additional context, watermarking embeds information directly into the content to prevent alteration.

4. Criminal Justice Administration – Deepfake:

First, any legal standard quickly becomes outdated due to the speedy progress of technology. Second, defining what and how technology should be used is critically important. Third, because these technologies are cross-border in nature, it may be very difficult to determine which regulations these technologies need to abide by; for this reason, they are typically registered in nations with laxer regulations. Fourth, when technology development and usage are not limited to a single nation, it is challenging to enforce laws. Fifth, there is often partiality in the obligations of the parties involved in deepfakes. Deepfakes can have extremely detrimental effects on people's reputations, produce false events or content that could lead to an incorrect conviction or influence legal actions, erode public confidence in institutions, endanger national security, or damage relations with other countries if used improperly by governments. Deepfakes can be extremely hazardous and damaging from the standpoint of criminal justice administration since they blur the lines between what is true and what is false. Judges, attorneys, and police officers may be duped by the presentation of manipulated images, sounds, or documents as evidence, "casting doubt on audio-visual evidence as an entire category of evidence." We can get an indication of the possible stance that the European Union may take on this issue from the proposal for a Regulation of the European Parliament and of the Council laying out coordinated laws on artificial intelligence (Artificial Intelligence Act). Article 1(d) lists the guidelines for AI systems that are used to "generate or manipulate image, audio, or video content" as part of its subject matter.

Article 3(1) defines artificial intelligence (AI) systems as "software that is developed with one or more of the techniques and approaches listed in Annexe I and can generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with, for a given set of human-defined objectives. The artificial intelligence method used by this suggested paradigm is risk-based and divides risk into four categories: intolerable, high, limited, and low. It is vital to provide a brief definition of each deepfake to better comprehend the degree of risk that they pose. Despite mentioning a three-tier risk-based strategy (unacceptable, high, and low or minimal), the European proposal also acknowledges a fourth level that is widely acknowledged: little risk. Article 52 of the proposal states that "certain" AI systems are obligated to maintain transparency. This means that while their use is permitted, users and providers are required to disclose to end users that they are engaging with artificial intelligence (AI) systems or material. This risk level covers a wide range of artificial intelligence (AI) systems, including those that interact with humans, recognise emotions, classify data based on biometrics, or create deepfakes. Determining whether deepfakes need to be declared or identified as such, or even if they are lawful, requires case-by-case consideration. In certain situations, such as those involving defamation, extortion, or child pornography, this kind of fake content may be considered a criminal violation of the basic rights of third parties. For others, though, it can just be a creative expression. It can be challenging to define the bounds of the freedoms of expression, the arts, and the sciences. Reducing the possibility of widespread manipulation through deepfakes could prove to be an even more formidable obstacle for legal measures to surmount. Technology-based content restriction or identification systems may prove to be beneficial for individuals, companies, and governments alike.

5. What can India do?

Deepfakes have three stages in their life cycle: creation, propagation, and detection. Regulating AI can help reduce the production of illegal or non-consensual deepfakes. Countries like China are addressing this kind of regulation by requiring deepfake technology suppliers to get the approval of people in their films, confirm users' identities, and provide them with a remedy. Mass public awareness efforts and potential legislation that would make the creation and distribution of deepfakes with malicious intent unlawful are two of Canada's strategies for preventing the harm that deepfakes can cause. Watermarking videos produced by AI is necessary for efficient identification and crediting. Watermarks provide a variety of functions by disclosing the source and owner of the information. By identifying the author or source of the work, they facilitate attribution, particularly when disseminated in various contexts. Moreover, watermarks that are visible serve as a warning that content can be traced back to its source and discourage unauthorised use. Watermarks also facilitate accountability by making it easier to enforce copyright and intellectual property laws for work created by AI by supplying proof of the original creator's rights. This may entail creating new techniques that can recognise deepfakes based on metadata, context, or other elements, in addition to employing increasingly complex algorithms. Since they are unable to stop the construction and initial distribution of deepfakes, the provisions of the current IT law might not be enough to address the issue. According to Siddharth Deb, manager of public policy at TQH Consulting, "criminal provisions under the IT Act and the IPC only partially address the harms which arise from deep fakes." "Policymakers need to find solutions that deal with this and lessen the psychological toll on victims." As the UK considers legislation that combines accountability and transparency through the labelling of deepfake photos and videos, India might work with other countries to investigate content labelling solutions or watermarking of AI-generated content.

CONCLUSION:

Enforceability is the main issue that the criminal justice system needs to resolve. The legal system in place is insufficient to combat deepfakes. In summary, the issue of applying current legal regulations to deepfakes can be summed up as follows. First, any legal standard quickly becomes outdated due to the speedy progress of technology. Second, defining what and how technology should be used is critically important. Third, because these technologies are cross-border in nature, it may be very difficult to determine which regulations these technologies need to abide by; for this reason, they are typically registered in nations with laxer regulations. Fourth, when technology development and usage are not limited to a single nation, it is challenging to enforce laws. Deepfake mitigation becomes crucial. A secure deepfake mitigation approach was suggested by us. Using a formal security verification method based on the Scyhter tool, we presented a security analysis of the suggested framework. It demonstrated the suggested framework's resistance to several types of attacks. We also talked about how deepfake events affect society and how to spot them. Lastly, we offered the suggested framework's actual implementation so that users could see how it operated in a real-world setting. We hope to enhance the suggested framework's capability in the future. Additionally, various techniques based on machine learning can be applied to the identification and mitigation of deepfake scenarios.

REFERENCES:

1. Abrams v. United States, 250 U.S. 616 (1919).
2. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, § rule3(1)(b).
3. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, § rule3(2)(b).
4. Jain, A. (2025). Deepfakes and misinformation: Legal remedies and legislative gaps. *Journal of Law & Social Development,* 3(2), 56–72. Verma, K. (2024). Digital deception: The impact of deepfakes on privacy rights. *Law, Society & Legal Review, 10*(1), 112–128.
5. Mohan, S., & Wadhwa, S. (2023). Deepfakes and shallow laws: Regulating distorted narratives in the political cyberspace. *Indian Journal of Law & Technology, 19*(2), 85–116.
6. Birrer, A., & Just, N. (2024). What we know and don't know about deepfakes: An investigation into the state of the research and regulatory landscape. *New Media & Society*.

***