DEEPFAKES AND THE LAW: A COMPREHENSIVE COMPARATIVE ANALYSIS OF INDIAN AND INTERNATIONAL LEGAL FRAMEWORKS

Ms. Suruchi, Research Scholar, Department of Law, Deen Dayal Upadhyaya Gorakhpur University, Civil Lines, Gorakhpur - 273009, Uttar Pradesh

ABSTRACT

Deepfake technology, which generates synthetic media using artificial intelligence, has created unprecedented challenges for judicial systems worldwide. Deepfakes undermine the credibility of digital content, raising concerns about consent, privacy, misinformation, defamation, and democratic discourse. This thesis examines deepfake regulation in India from a legal and ethical perspective, using a global comparison. The study focuses on Indian law, analyzing current legislative provisions in the Bhartiya Nayaya Sanhita, 2023 and Information Technology Act of 2000, as well as constitutional guarantees including free speech and privacy. This assessment evaluates how effectively laws address the complex liabilities associated with deepfakes, including impersonation, non-consensual pornography, political disinformation, and reputational damage. The study reveals significant statutory and interpretation gaps, limiting alternatives for deepfake victims under existing systems. This article compares regulatory frameworks in the United States, European Union, United Kingdom, South Korea, Australia, Japan and China to understand India's worldwide position. It emphasizes legal advances and best practices. The goal of laws like criminalization, platform responsibility, transparency requirements, and judicial decisions is to strike a balance between fundamental freedoms and technology progress. According to the article, India needs a customized legal structure to handle the dangers posed by deepfakes. Clear statutory definitions, procedural safeguards, victim-oriented remedies, and intermediary responsibility should all be included in the framework. Pursuant to the report, maintaining the rule of law and defending democratic principles in the digital era necessitates a proactive, rights-based, and technologically astute legal strategy.

Keywords: Deepfakes, Artificial Intelligence, Generative Adversarial Networks, Synthetic Media, AI Regulation, Information Technology Act, Digital Personal Data Protection Act, Cybercrime, Privacy, Non-Consensual Pornography, Political Misinformation, Financial Fraud, AI Content

Labeling, Platform Accountability, Digital Literacy, India, International Legal Frameworks, Rights-Based Regulation.

INTRODUCTION

Emerging technologies have the potential to shape humanity's future by enabling better outcomes. Artificial intelligence has led to the development of numerous new technologies. The term "deepfake," which combines the terms "deep learning" and "fake," was first used in 2017 by a Reddit user who created a community for discussing AI-generated modified films, particularly explicit face-swapped footage of celebrities. Deepfakes are classified as synthetic media and belong to the family tree. AI has been widely used since its conception and continues to evolve through its various applications. AI can analyze and adjust data sets, both internal and external, to meet prescribed goals and tasks. AI has transformed various industries, including healthcare, retail logistics, and self-driving cars.

Deep-fake technology is a new digital technology that has gained popularity alongside others. The key sections and definition will be explained later, but here is a basic summary: Deepfake uses artificial intelligence (AI) to create a synthetically augmented video of a real person, complete with audio-visual signals that cause them to speak or act in ways that are not realistic.

Machine learning algorithms can re-generate any human part on screen, including the face, body, and other visual features that appear realistic but are not. The problem profile highlights the need for a framework to govern the use of emerging technologies, including deepfakes, both domestically and internationally. Indian involvement or interference may be necessary to address this issue. Three distinct types are used by researchers to classify deepfake content: Video editing comes in three flavors: puppet-master, lip-sync, and face swap. While Lip-sync adjusts a source video to match an audio recording, Face Swap automatically swaps out a person's face in a video. A puppet-master is a performer who uses their head, eyes, and facial expressions to mimic the actions of their puppet. The infinite inventiveness of technology is exemplified by deepfake films of Queen Elizabeth, Donald Trump, Nancy Pelosi, Russian President Vladimir Putin, Ukrainian President Volodymyr Zelenskyy, Malaysia's Minister for Economic Affairs Azmin Ali, and Barack Obama. The Deepfakes Accountability Act explains how Deepfake films or images can cause harm across various domains. The 9th Act criminalizes using advanced technology to create false personas with the intention of promoting sexual activity, inciting violence, obstructing official action, engaging in deception,

influencing domestic policy, or tampering with elections.

Deepfakes offer a significant threat due to their high accuracy, ease of production, and harmful influence on viewers. Untrained humans and AI systems consider it more challenging to distinguish between real and false videos as deepfake film quality increases. Deepfake detection is difficult because to the constant threat. Women, both known and unknown, are particularly vulnerable and exploited by pornographic deepfakes. A deepfake vengeance porn film with an Indian Muslim female investigative journalist was widely released in April 2018. As a consequence, her private information was leaked, she received inappropriate offers, and she received violent threats on her private phone and social media accounts. In India, deepfakes are a relatively new issue, and there is currently no explicit regulation to address their effects. Existing laws, however, offers potential legal recourse. Forgery, defamation, criminal intimidation, breach of peace, communal disharmony, sedition, computer-related offenses (e.g. transmission or publication of pornographic or sexually explicit images), identity theft, cheating by personation, violation of privacy, and infringement of intellectual property rights are examples of criminal penalties. Indian legislation primarily penalizes creators of deepfake videos through civil and criminal sanctions. Disseminators of such videos have complete immunity and are not held accountable for their actions.

HISTORICAL EVOLUTION OF DEEPFAKES

Deepfake technology emerged from early studies with artificial intelligence and image manipulation, quickly maturing into a sophisticated approach for producing hyper-realistic fake media. The term "deepfake" was coined in 2017, however the underlying concepts stretch back far earlier. This technique began in the 1990s, with breakthroughs in video and picture manipulation. However, Ian Goodfellow and colleagues made a huge breakthrough in 2014 by introducing Generative Adversarial Networks (GANs). GANs are made up of two neural networks: a generator that generates synthetic data and a discriminator that evaluates its validity. By striving to produce better images, they enable increasingly realistic fake visuals. Deepfakes became better as GANs got better. Multilayer convolutional neural networks were incorporated by researchers around 2016–2017 to enhance image synthesis and identification, permitting for more lifelike face swaps and video editing. In the same year, Nvidia released algorithms for training that made it possible for GANs to produce images with increasing resolution in stages, making it more challenging to spot fakes. The public became aware of

deepfakes mostly through online discussions and social media. A Reddit user going by the handle "deepfakes" made the production and dissemination of altered videos, with particular emphasis on non-consensual sexual content starring celebrities, more common in late 2017. Issues regarding potential abuse, such as political influence, disinformation, fraud, and privacy issues, were raised by this. Recent breakthroughs have extended deepfake capabilities beyond images and videos to include speech synthesis and real-time modifications, posing additional ethical and regulatory problems. This historical progression illustrates deepfake technology as a double-edged sword, providing creative opportunities while also posing major threats to privacy, trust, and democratic processes. Understanding its growth informs continuing efforts in detection, law, and media literacy to limit potential effects, responsible

Types of Deepfakes:

1. Face swapping: The most well-known type, face swaps overlay a subject's face over another person's body while in motion. Neural networks excel at tracking expressions and matching them frame by frame to create realistic illusions. Some of these deepfake movies are malicious forgeries that could negatively impact reputations, while others are humorous memes. Even the most observant viewers without sophisticated detection skills can be confounded by high-fidelity detail.

2. Lip-Syncing and Audio Overlays: Lip-sync fakes, often known as 'puppeteering,' use mouth movements to match synthetic or modified audio. The words are never addressed to a speaker, but they appear to be. With voice cloning, the 'face' in the video can convincingly perform complete scripts.

3. Voice- only Cloning: Audio deepfakes rely entirely on the imitation of the AI speech without visuals. They are used by scammers in phone scams, such as impersonating an executive to request urgent wire transfers. Recording "celebrity cameo" voiceovers for marketing purposes is something that some individuals do. Since this type of deepfake lacks visual indicators and necessitates in-depth spectrum analysis or ambiguous context triggers, detection is challenging.

4. Full-Body Reenactment: Generative models may record an actor's whole posture, movement, and gestures and map them to another person. The ultimate result is a subject who appears to be dancing, playing sports, or completing duties that they have never done.

Film or virtual reality experiences require full-body illusions. With deepfake cybersecurity, however, the potential for producing 'alibi videos' or counterfeit proof is of particular concern.

5. Text-Based Conversational Clones: While not as well known as deepfakes, generative text systems mimic a person's writing style or dialog. Cybercriminals create new message threads imitating the user's language and writing style.A multi-level fake or even a full deepfake character can be created by adding the voice or image to the illusion. Furthermore to image fraud, text-based generative AI is anticipated to be used in social engineering techniques through chat platforms as it improves in complexity.

How deepfakes are created?

1. Face-Swap Apps: There are a number of consumer programs available that allow beginners to easily generate face swaps from their phones or computers. The application submits two movies (one source and one target) to automate training and combining. However, these applications can be used to create false identities if they are utilized maliciously. Both substantial deepfake misuse and lighthearted fun are encouraged by democracy.

2. GAN Frameworks and Open-Source Code: Advanced findings are available through frameworks such as TensorFlow or PyTorch, which include specific repositories for face and voice forging. Network designs can be altered, training settings can be changed, and multiple data sets can be integrated by tinkerers who understand network architecture, training parameters, and data combinations. The best deepfakes can be generated with this method, but it demands more coding knowledge and hardware (GPU). This enables you to fine-tune illusions beyond the pre-made presets, which increases the bar for deception significantly.

3. Audio-Based Illusions: Creators of audio-based illusions use voice synthesis scripts in conjunction with lip sync modules to achieve realistic mouth movements. To generate fresh lines that mimic the target's accent or demeanor, the algorithm can be trained using voice samples. Aligning the lips guarantees that every uttered phoneme is reflected in the visuals. These "deepfake lip-sync combos" can produce "talking head" illusions which appear remarkably realistic.

4. Cloud-Based Rendering Services: Some commercial deepfake providers or AI tool suppliers can handle extensive model training on distant servers. Users just provide data sets or script parameters and wait for the final results. Due to the removal of local GPU constraints complex or big illusions can now operate on reliable systems. However, it makes it feasible to produce sophisticated forgeries quickly, which raises questions concerning deepfake cybersecurity.

5. Manual Overlays and Hybrid Editing: After generating a neural net-based face map, creators use software such as Adobe After Effects to manually enhance frames. For as few artefact transitions as possible, they employ shallowfake splices, alter lighting, or remove boundary artifacts. Skilled post-production in combination with AI-generated content is almost ideal. The outcome is a complete fake that can easily place a phony topic anywhere, from sarcastic impersonations to humorous comedies.

CHALLENGES POSED BY THE UNFAIR USE OF THE TECHNOLOGY

Deepfakes pose a threat to individuals' privacy and liberty, which are protected by constitutions worldwide. Freedom of speech and expression is a guiding principle in various constitutions, ensuring individual autonomy for citizens. As previously stated, deepfakes are morally ambiguous and pose ethical challenges. To address morally ambiguous circumstances, a constitutional right can be a viable answer. It is important to sub-categorize the negative impact under this investigation. Digital injury encompasses physical, mental, and economic harm, as well as reputational harm, which is covered by other legal rights such as defamation. The creator's intent has a vital role in evaluating moral and societal harms. Deepfakes, whether intended or not, continue to inflict harm even if they exceed expectations when shared on social media. Women's rights must be prioritized in the fight against deepfake abuse caused by the proliferation of pornographic content, which affects celebrities and women in general. Deepfakes tend to create non-consensual erotic content aimed at women rather than political satire or misinformation.

In addition to regulatory action, further measures are needed to address injury to other rights under various statutes. The punishment or fine is automatically covered by the statute that applies to the act or harm committed. Current legislation needs to widen its scope due to the lack of date protection regulations.Regarding any sort of media that is shared on platforms. To protect individuals' rights, it is important to take responsibility for the platforms where such

content is found, implement corporate policies to raise awareness, and develop better detection technologies for suspicious data. The criminal law section will examine the relationship between deepfakes and this area of legislation, as well as the impact they have had on the evidence landscape in court. The review will be based on relevant case laws, both domestic and global, and will include recommendations for implementing detection systems in courts. The IP concerns mostly focus around copyright infringement and whether AI-based deepfake technology should be protected by copyright, especially for machine-generated content. The Copyright Act of 1957 in India provides protection for many aspects, but it may not adequately address developing technologies like deepfakes. Additionally, the question of whether creators should have a copyright remains unanswered. The original deployment of deepfakes in mainstream media targeted women. Deepfakes pose a risk to public figures who have videos of themselves speaking on the internet. Because deepfakes do not allow for interaction between the abuser and victim, a third party, such as a regulatory body, is necessary to balance the gap.

In the entertainment industry, deepfakes have transformed filmmaking and digital content development. Hollywood movies now use deepfake techniques to de-age actors or construct lifelike digital doubles, as shown in Martin Scorsese's "The Irishman," allowing for seamless storytelling without physical limits. Hyper-realistic avatars based on deepfake-generated facial and behavioral models improve immersion and personalization in video games and virtual reality. Deepfakes are often used in satire and parody to provide surprisingly realistic humor and social commentary. Deepfakes offer novel educational and accessibility experiences, like the ability to create synthetic voices for individuals with speech impairments to improve communication and social interaction, or the ability to digitally bring back historical figures (like Agatha Christie) to teach lessons and lectures. Users can create personalized digital personas for identity representation and self-expression with AI-generated avatars. Deepfakes enable the development of highly customized and targeted advertising strategies. In order to improve consumer engagement and sales conversion, retailers use deepfake-generated models to power virtual trial rooms where shoppers can virtually try on clothing and accessories. By distributing AI-generated content at scale, influencers and celebrities can expand their audience and enhance fan engagement. Deepfakes have played a critical role in digital forensics and public safety, assisting in crime scene reconstructions using synthesized media and integrating varied sources such as surveillance footage and forensic reports to create cohesive virtual narratives.

However, deepfake technology raises severe cybersecurity and ethical concerns. The capacity to convincingly imitate individuals opens the door for a range of undesirable activities:

- Identity theft and fraud: Criminals employ deepfakes to establish falsified credentials, synthetic identities, or mimic CEOs in order to get illegal access to financial resources or sensitive information. Deepfake speech fraud, for example, has been utilized in millions of dollars worth of CEO scams around the world.

- False information and political deception: Deepfake films of political leaders and public personalities circulated during election seasons are intended to affect public opinion, destroy trust, or incite social unrest. The dissemination of falsified speeches or occurrences can seriously jeopardize democratic processes.

- Non-consensual personal content: The development and transmission of deepfake pornography, which frequently targets women without consent, is a serious invasion of privacy, exacerbated by the hyper-realistic nature of such content, resulting in significant mental health consequences, harassment, and reputational injury.

- Blackmail and harassment: Deepfake content is increasingly being used for extortion and harassment, with legitimate media being manipulated to create compromising scenarios in order to intimidate or silence victims.

- Financial market manipulation: Forged movies or audio can be used to influence stock prices by attributing phony statements to corporate officials or broadcasting bogus product announcements.

- Social engineering attacks: Deepfake sounds or videos impersonating trusted individuals are used in social engineering attacks to trick victims into disclosing sensitive information or carrying out fraudulent activities.

The duality of deepfake technology requires balanced governance. While the creative and educational benefits are significant, strict controls must be designed and enforced to reduce hazards. Detection systems are evolving, but the rapid advancement and availability of deepfake production tools make enforcement difficult. In conclusion, a thorough analysis of

the socio-legal ramifications of deepfakes is necessary given their increasing usage. To maximize potential while reducing misuse, appropriate legislative frameworks, technological detection tools, and public awareness are needed. This ensures that deepfake technology benefits society without endangering security, privacy, or confidence. Since its debut in 2025, deepfake technology has expanded quickly, primarily due to developments in Generative Adversarial Networks (GANs), which produce more accurate and realistic artificial media that is hyper-realistic. With digitally created avatars and immersive experiences, almost two-thirds of evaluated deepfake material already closely resembles genuine videos, creating new opportunities in virtual reality, education, and entertainment.Concurrently, detection approaches have evolved, combining AI-based neural anomaly detection, metadata analysis, and adversarial training to achieve annual accuracy gains of over 40%. However, even static detection systems struggle against ever-changing fakes, necessitating the change to adaptive, constantly updated defenses.The ubiquitous availability of open-source AI tools democratizes deepfake production, increasing creativity but also amplifying the potential for abuse. The increasing threat of vocal cloning with emotional complexity and regional accents has led to an increase in the intricacy of phishing and social engineering methods. The market for deepfake detection is anticipated to reach a valuation of over $3.5 billion by 2025, demonstrating the growing demand for effective cybersecurity solutions. Because of these trends' dual nature— their unprecedented creative potential and growing risks—technological, governmental, and educational strategies must be coupled to preserve trust in digital communications while fostering responsible innovation.

THE INDIAN LEGAL FRAMEWORKS ON DEEPFAKES

Deepfakes—manipulated audio, video, and image information produced using sophisticated artificial intelligence techniques that have the power to mislead, slander, and disrupt society—have become a rising problem for India in recent years. Policymakers and regulators in India have worked to create a strong legal framework to combat the serious hazards of false information, identity theft, political manipulation, and privacy violations carried on by synthetic media. A historic regulatory endeavor, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2025, require the unambiguous labeling and traceability of information produced by artificial intelligence. This evolving legal architecture expands on existing laws such as the Information Technology Act of 2000 and the Digital Personal Data Protection Act of 2023, while assigning new responsibilities to

intermediaries and social media platforms to quickly identify, restrict, and remove illegal deepfake content. India's legal response illustrates a delicate balancing act between supporting AI innovation and safeguarding citizen rights, public trust, and democratic procedures in an increasingly digital world. This introduction lays the groundwork for a thorough examination of the Indian legal frameworks governing deepfakes, emphasizing recent legislative changes, enforcement methods, and current issues.

Existing statutory provisions applicable (IT Act, BNS, IPC, etc.)

India's legal framework to combat deepfakes is mostly based on existing statutes such as the Information Technology Act of 2000 (IT Act), the Bharatiya Nyaya Sanhita of 2023 (BNS), and the Digital Personal Data Protection Act of 2023 (DPDP Act). TThe IT Act has provisions that address privacy violations (Section 66E), identity theft (Section 66C), impersonation (Section 66D), and the dissemination of unlawful or pornographic material (Sections 67, 67A). Additionally, it mandates due diligence from intermediaries (Section 79) and permits the government to impose blocking orders (Section 69A). Additional intermediary requirements, including the prompt removal of illegal content, the creation of grievance redressal procedures, and the requirement for transparency when handling synthetic or deepfake information, are outlined in the 2021 IT Rules, which were amended in 2022 and 2023. The DPDP Act protects personal data processing by prioritizing user permission and security, whereas BNS Section 353 criminalizes misinformation causing public mischief, which can include deepfake-related harms, and has a greater law enforcement scope. India's institutional systems, such as the Indian Cyber Crime Coordination Centre (I4C), CERT-In, and Grievance Appellate Committees (GACs), supplement legal requirements by easing coordination, reporting, and enforcement of deepfake-related cybercrimes. Through the implementation of technical measures like mandatory watermarking, metadata embedding, and visible labels on AI-generated content, the proposed amendments to the IT Rules aim to make intermediaries more proactive. A 'Techno-Legal' crackdown on deepfakes has also been alluded to by the government, highlighting the need for clear regulations to address the emerging threats. However, critics claim that, despite these attempts, there is a lack of particular deepfake legislation, which complicates enforcement, accountability, and the balance of free expression and security.

Regulatory frameworks and government initiatives

The legislative frameworks and government activities addressing deepfakes in India represent

a changing landscape that aims to reduce misuse while balancing innovation and digital rights. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, are at the heart of this, requiring social media intermediaries and digital media firms to delete illegal content as soon as possible. The 2025 draft amendments, which require content hosts and intermediaries to label or watermark content that is synthetically generated, including deepfakes, are the most notable change to these principles. In an effort to promote transparency and responsibility in the digital ecosystem, the amendments mandate that synthetic content have labels that span at least 10% of the visible surface or 10% of the audio duration, along with integrated information that permits traceability. Furthermore, operational enforcement and capacity building are significant functions of government organizations like the Indian Cyber Crime Coordination Centre (I4C), Computer Emergency Response Team-India (CERT-In), and the Ministry of Electronics and Information Technology (MeitY). CERT-In, for example, has issued advisories with cybersecurity rules for detecting and reporting deepfake risks. Citizens can easily report deepfake-related offenses through the National Cyber Crime Reporting Portal, allowing for timely intervention. The formation of Grievance Appellate Committees increases redress processes and holds intermediaries accountable. India's multi-stakeholder approach also involves awareness campaigns and public outreach activities to educate citizens about the risks of deepfakes.

Critical analysis of legal gaps and challenges in India

India's legal approach to deepfakes indicates substantial gaps and obstacles that impede effective prevention and enforcement against the abuse of this rapidly growing technology. The absence of a clear legal definition of "deepfakes" or "synthetic media" in Indian statutes is a significant issue that makes legal interpretation and responsibility more difficult. Current laws, like the Bharatiya Nyaya Sanhita of 2023 (BNS) and the Information Technology Act of 2000 (IT Act), deal with similar crimes like identity theft, impersonation, and disinformation, but they don't specifically address the complexities of artificial intelligence-generated synthetic content. This results in disjointed solutions that don't solve deepfake-specific problems like malicious fabrication, digital permission, and non-consensual intimate content. Furthermore, the Indian Penal Code does not specifically address the malicious development and distribution of these synthetic media forms, reducing prosecuting incentives. Platform liability and intermediary responsibilities present yet another significant obstacle. Despite imposing due diligence requirements and requiring social media platforms to promptly remove illegal

content, the IT Rules, 2021, together with their 2025 modifications, face enforcement hurdles and constitutional challenges pertaining to freedom of expression. Mandates for transparency, like watermarking or naming deepfake information, are very new and don't have strong compliance systems. Law enforcement has a hard time catching the persons who make deepfakes because of anonymization, encryption, and inexperience. People are more vulnerable to scams and misleading information when they lack digital literacy and public awareness about deepfakes. A comprehensive deepfake law that balances innovation, privacy, and security without compromising free speech is desperately needed in India, as evidenced by the stark contrast between the glacial pace of legislative reform and the rapid advancements in technology. India's current legal framework addresses deepfake issues in a basic but inadequate manner, necessitating urgent changes to create clear statutory definitions, strengthen intermediary accountability, enhance forensic capabilities, and give victims targeted legal remedies. Without these, deepfake abuse will continue to exploit legislative loopholes and institutional flaws, posing greater threats to privacy, democracy, and public trust.

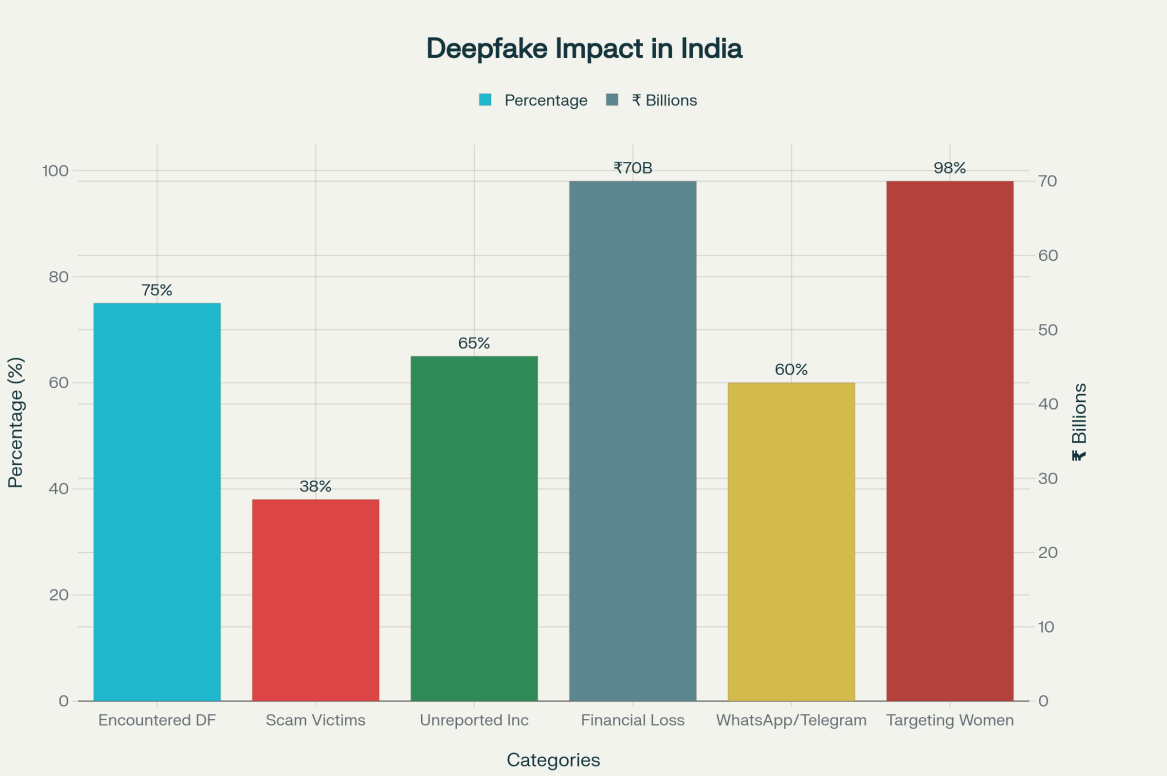
Empirical Landscape and Data in India

India is experiencing an increasing epidemic of deepfake influence, with polls finding worrying levels of exposure, vulnerability, and economic impact. According to a 2025 survey, almost 75% of Indians with internet connection have experienced some type of deepfake content in the previous year, indicating a broad infiltration of falsified media into daily life. Among these, 38% reported falling victim to deepfake scams, in which audio or video impersonations were utilized for fraud, extortion, or misleading information. Due to a lack of understanding, a significant portion of cyber occurrences involving deepfakes—more than 65%—go undetected, impeding efficient law enforcement and response. According to estimates, crimes associated to deepfakes, including financial fraud, defamation, and reputation attacks, might cost India up to ₹70,000 crore (about $8.4 billion) in 2025 alone.

WhatsApp and Telegram are the most common platforms for spreading deepfakes, causing over 60% of instances, according to a survey conducted throughout major cities. The attacks vary from celebrity defamation and non-consensual pornography, which primarily affects women (98% of deepfake porn content targets female victims), to political disinformation with phony recordings of politicians. The impact is both societal and infrastructural; India's financial industry sees an increase in deepfake scams targeting banking and investing platforms, with

losses exceeding 550% since 2019. These statistics highlight the critical need for targeted policy, awareness initiatives, and technological safeguards to address the growing threat landscape.

CASE STUDIES OF DEEPFAKES INCIDENTS IN INDIA-



SURVEY DATA ON DEEPFAKE PREVALENCE AND IMPACT IN INDIA (2025)

Deepfake technology has emerged as a major challenge in India, affecting politics, entertainment, and security. As deepfakes—realistic AI-generated films and audio manipulations—have grown in popularity, India has seen an increase in events that highlight both the potential for abuse and the urgent need for effective defenses. Deepfakes were used in recent elections to distort political narratives and affect public opinion, reaching millions of voters while confounding the verification of real material. High-profile examples involving Indian celebrities who were subjected to unlawful deepfake movies highlight the harm to personal privacy and reputation. Furthermore, deepfakes have been used in financial schemes and misinformation operations, exposing weaknesses in national cybersecurity systems.

1. **Deepfake Political Misinformation Targeting Bollywood Celebrities: The 2024 Aamir**

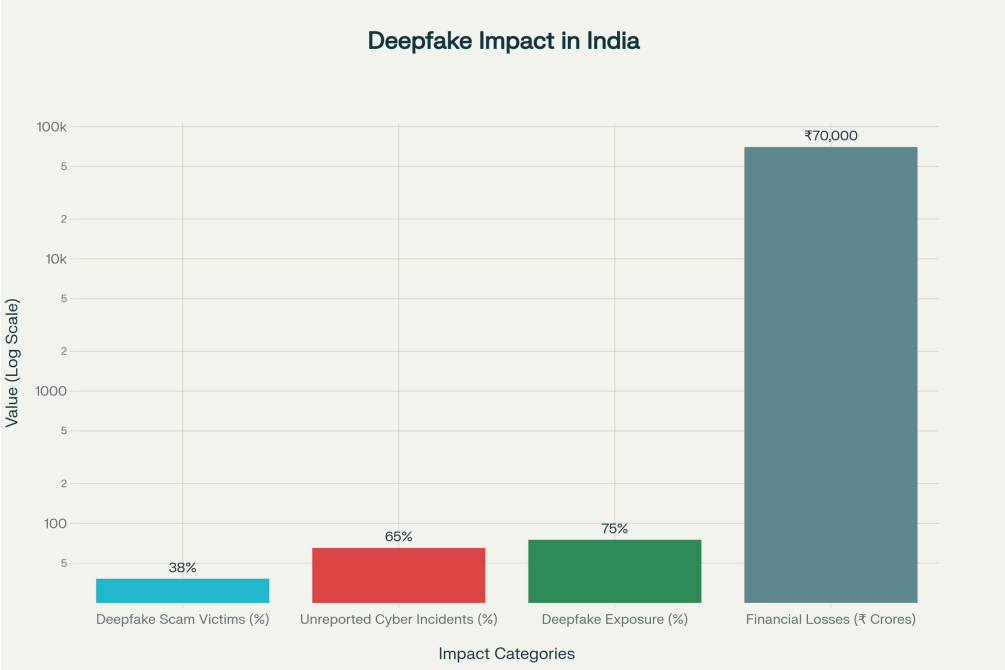
Khan and Ranveer Singh Deepfake Controversy: India has seen many high-profile deepfake occurrences, highlighting the mounting issues faced by AI-generated synthetic media. One noteworthy case included Bollywood actors Aamir Khan and Ranveer Singh during the 2024 general elections, when deepfake videos of them criticizing the Prime Minister went viral on social media. Both actors filed police complaints, underlining deepfakes' deceitful ability to sway public opinion and disrupt democratic processes. Another well-known victim was veteran performer Hema Malini, who expressed worry in Parliament about the use of AI-generated content to damage reputations and create emotional distress.

2. **Celebrity Exploitation and Non-Consensual Pornography:** Rashmika Mandanna Target (2023): A prominent case of celebrity exploitation occurred in 2023, when actress Rashmika Mandanna's face was morphed onto pornographic movies without her permission, causing uproar and legal action. In October 2025, Telugu superstar Chiranjeevi made a complaint with the Hyderabad Cyber Crime Police regarding deepfake videos depicting him in indecent actions, prompting a thorough investigation under Indian cyber laws. These examples demonstrate how deepfakes can breach privacy, dignity, and reputation, forcing victims to overcome significant legal and technological barriers in order to seek justice.

3. **Financial Fraud Deepfake Scams: Jamtara 2.0 Cybercrime Rings:** Financial fraud employing deepfakes has grown into a sophisticated threat known as "Jamtara 2.0," in which cybercriminals use voice cloning and video manipulation to impersonate corporate executives and government authorities. Cases from 2023 and 2024 show schemes in which AI-generated voices of finance ministers or CEOs persuaded regulators to approve false transactions, resulting in multi-crore losses. A 73-year-old man was tricked into sending ₹40,000 in one confirmed scam after he received a video call from a deepfake imitation of his former coworker. This draws attention to the dangers that regular people confront. With a 550% increase in cybercrime cases since 2019, deepfake fraud is increasingly focusing on digital banking, underscoring the pressing need for technological and legislative solutions. These case studies highlight serious implications for regulation, enforcement, and public awareness by demonstrating the widespread, dynamic, and lethal nature of deepfake misuse in India's political, entertainment, and financial domains.

Deepfake scams impersonating high-profile figures such as Finance Minister Nirmala Sitharaman and Google CEO Sundar Pichai have been reported, duping citizens into fraudulent investment schemes and cryptocurrency scams. According to estimates, 38% of Indian internet users will be victims of deepfake schemes in 2025, resulting in millions of rupees lost to sophisticated fraud.

SOCIETAL AND ECONOMIC IMPLICATIONS-



Deepfake technology has major societal and economic ramifications for India, highlighting a fast evolving danger scenario. Deepfakes have destroyed public trust, amplified disinformation, and increased exposure to harassment and slander. Approximately 75% of Indian internet users have come across deepfake content, however 65% of cyber incidents are unreported, exposing knowledge gaps and systematic enforcement issues. With 38% of consumers reporting direct scams that use deepfake audio or video for identity fraud or manipulation, victimization affects a wide range of categories. These online lies have caused reputational damage, disturbed political debate, and complicated socio-political communication. The cost of deception facilitated by deepfakes is enormous from an economic standpoint. By 2025, deepfake crimes are predicted to cost India ₹70,000 crore (Seventy thousand crores), or around $8.4 billion, and include financial fraud, identity theft, and extortion schemes that target both individuals and organizations. Deepfake-driven scams have increased more than fivefold since 2019, hurting industries ranging from banking to

entertainment, depleting resources and weakening trust in digital transactions and platforms. This economic impact poses a substantial threat to national cybersecurity and financial stability, necessitating immediate legislative reforms and technological countermeasures to combat the spread and consequences of synthetic media.

SURVEY DATA ON DEEPFAKES IMPACT IN INDIA ON SCOCEITAL AND ECONOMIC EFFECTS, (2025)

This data emphasizes the need for comprehensive legislative frameworks, strong enforcement, better detection technologies, and widespread public education to manage the various hazards that deepfakes bring to Indian society and the economy.

COMPARATIVE ANALYSIS OF INTERNATIONAL LEGAL FRAMEWORK-

The comparative research of foreign legal frameworks for deepfake regulation sheds light on how various jurisdictions approach the multiple difficulties provided by synthetic media. As deepfake technology rapidly grows, foreign authorities have taken a variety of approaches, influenced by their legal traditions, technological settings, and societal norms. Prominent jurisdictions include the United States, the European Union, China, and the United Kingdom, which have various regulatory philosophies ranging from decentralized state-level regulations to comprehensive, risk-based AI plans and centralized oversight. This analysis reveals differences in enforcement mechanisms, constitutional interpretations, and user protections, as well as convergences such as the emphasis on transparency, mandatory labeling of synthetic content, and platform accountability. Understanding these international models provides valuable lessons for developing balanced, context-sensitive policies, particularly for countries like India that want to protect citizens from deepfake harms while also encouraging technological innovation and protecting fundamental rights.

1. United States: State and federal deepfake laws

The United States' approach to deepfake regulation is characterized by a developing mosaic of state-level rules and potential federal legislation, each addressing a particular damaging feature of synthetic media. At the state level, California leads with several key laws, including AB 730 (2019), which criminalizes the distribution of deepfake videos depicting sexually explicit content without consent and prohibits manipulated political videos within 60 days of elections that are classified as materially deceptive and intended to harm candidates. Another crucial

provision is AB 602, which allows victims of deepfake pornography to file civil claims. Other jurisdictions, such as Texas and New York, have passed laws criminalizing the non-consensual creation and distribution of deepfakes, with a focus on intimate photographs and election propaganda. The DEEPFAKES Accountability Act, a major proposed federal bill, seeks to safeguard consumers from fraud by requiring clear, visible disclosures of AI-generated or synthetically altered information. The legal toolset also includes enforcement of basic consumer protection legislation employed by the Federal Trade Commission (FTC) as well as fraud statutes pursued by the Department of Justice (DOJ) in deepfake-related frauds.

Importantly, constitutional free speech protections have influenced judicial decisions, such as the August 2025 federal court ruling that invalidated California's AB 2655, which sought to require platform takedown of materially deceptive election content, highlighting the delicate balance between regulation and First Amendment rights. Thus, US regulation combines targeted criminal sanctions and civil remedies, such as California's AB 730 and AB 602, with proposed federal transparency regulations like the DEEPFAKES Accountability Act, reinforced by broad enforcement of consumer protection laws. This varied legislative landscape reflects both proactive state-level innovation and complicated constitutional considerations governing deepfake media.

2. European Union: AI Act and transparency requirements

The European Union has the most extensive and organized regulatory approach addressing deepfake technology with its landmark Artificial Intelligence Act (AI Act), which is poised to set international norms for AI governance. The AI Act, which will be implemented in stages starting in 2024, establishes a risk-based regulatory framework that classifies AI applications based on their potential harm to basic rights and societal interests. Deepfakes are expressly addressed in clauses that require transparency, accountability, and user protection, reflecting the EU's precautionary attitude to AI-related dangers. A crucial feature of the AI Act is its transparency requirements, which require producers and deployers of AI systems that generate deepfake or synthetically edited information to clearly disclose to consumers that the content was artificially generated or altered. This disclosure must be conspicuous, understandable, and easily available, allowing consumers to discern between synthetic and actual content. The Act requires required watermarking or technical labeling of deepfakes to promote traceability and accountability, thereby discouraging misuse and increasing digital literacy. Furthermore, the

AI Act imposes stringent governance requirements on paperwork, compliance evaluations, and human monitoring for high-risk AI applications such as synthetic media designed for political impact, deepfake pornography, and security-sensitive situations. The AI Act allows enforcement organizations to levy hefty fines for noncompliance, incentivizing platforms and developers to incorporate detection and preventive methods into their AI ecosystems. The EU intends to encourage trust in AI by unifying transparency norms across member states, combating misinformation and protecting individual rights. The Act's ethical and legal improvements put the EU at the vanguard of worldwide attempts to regulate deepfakes, balancing technological innovation with strong safeguards for democracy, privacy, and free speech in the digital era. The EU AI Act's emphasis on mandated disclosures, watermarking, risk-based governance, and enforceable accountability creates a comprehensive framework for controlling deepfake hazards, acting as a significant point of reference for other jurisdictions considering synthetic media regulation.

3. China: Deep Synthesis Provisions and mandatory labeling

China has established a tight regulatory framework for deepfake technology with its "Provisions on the Administration of Deep Synthesis Internet Information Services" (also known as the Deep Synthesis Provisions), which went into force in January 2023. These rules address artificial intelligence-generated synthetic media by emphasizing openness, accountability, and control in order to prevent misuse and encourage innovation. A key component of the system is the obligatory labeling and digital watermarking of deepfake content, which requires operators to properly designate all artificially generated or altered media to distinguish it from genuine content. To guarantee traceability and public trust, these labels must be prominently displayed—covering at least 10% of the visual or auditory material. The restrictions also compel deep synthesis service providers to register with Chinese authorities, keep track of AI models, and apply strict content moderation procedures to prevent the spread of unlawful or harmful synthetic content. Real-name registration is required for users developing or distributing deepfakes, and violations result in severe penalties such as fines and criminal charges, showing China's unwavering stance against malevolent synthetic media. Unlike the European Union's AI Act, which takes a more risk-based and technology-neutral approach with some exclusions for artistic use, China's regulations are severe, centralized, and unconditional, with a focus on state control and social stability. Both frameworks seek to protect democratic integrity and individual rights, but China's approach

requires stricter labeling, user authentication, and proactive monitoring throughout the AI-driven content lifecycle.

4. United Kingdom: Online Safety Act and criminal provisions

The United Kingdom has taken a proactive and thorough approach to regulating deepfakes, enacting the Online Safety Act 2023 alongside modifications to the Criminal Justice Bill that specifically outlaw the creation and dissemination of sexually explicit deepfakes. The Online Safety Act criminalizes the sharing and threatening to distribute non-consensual intimate photographs, including deepfakes, and goes into effect on January 31, 2024. This act added Section 66B to the Sexual Offences Act 2003, which imposes severe penalties on people who transmit such content without consent. In addition, the Criminal Justice Bill (amended in 2024) makes it illegal to create intimate images or films using computer graphics or any digital technology with the intent to cause distress, fear, or humiliation, a ground-breaking development that extends protections beyond distribution to the act of creation itself. The UK's regulatory structure strikes a compromise between strong criminal consequences and respect for free expression and technological neutrality. The Act avoids specifically identifying "deepfakes" but tackles intentionally created synthetic content using language such as "using computer graphics or any other digital technology." Beyond intimate picture offenses, the Online Safety Act prohibits fraudulent communications, sending threatening messages, and inciting self-harm, giving a comprehensive shield against a variety of online ills amplified by AI-generated content. Ofcom is in charge of enforcement, and it has the authority to levy substantial fines on platforms that fail to comply with content moderation regulations.Though the UK lacks a single deepfake law, this set of legal tools, combined with ongoing consultations on AI regulation and copyright for synthetic media, positions Britain as a leader in comprehensively addressing AI-facilitated abuse while encouraging continued technological innovation within a secure legal framework.

5. South Korea, Australia, Japan: Selected legislative models

South Korea, Australia, and Japan each have unique legislative approaches to the difficulties brought by deepfake technology, reflecting their respective socio-legal settings and technological agendas. In South Korea, the legal framework governing deepfakes is largely centered on privacy and image protection. The Act on Special Cases Concerning the Punishment of Sexual Crimes and the Personal Information Protection Act (PIPA) both address

deepfake pornography, establishing severe penalties for the creation and distribution of non-consensual synthetic explicit media. South Korea distinguishes itself through proactive enforcement efforts and digital literacy campaigns that educate citizens on how to identify and report deepfake violations. Furthermore, constitutional safeguards balancing expression and privacy are critical in developing legislation that seeks to limit harm while safeguarding essential freedoms. Australia has taken a sophisticated approach, incorporating deepfake issues into existing legislation addressing image-based abuse, cybercrime, and defamation. The Enhancing Online Safety Act of 2015 was updated to add provisions against intimate image misuse, including AI-generated deepfake content. The eSafety Commissioner is in charge of enforcement, which includes reviewing content removal requests and imposing penalties on businesses that facilitate harm. Australia prioritizes victim-centered solutions and community engagement, combining legal sanctions with educational outreach and technology assistance. In response to growing public concern, the country is considering legal revisions aimed at combating misinformation and political deepfakes. Japan's approach is more emergent, with deepfake regulation incorporated within larger legislative revisions tackling AI ethics and cybercrime. The Act on the Regulation of the Transmission of Specified Electronic Mail and the Act on the Protection of Personal Information offer some legal protection against unauthorized synthetic media use. In 2023, Japan's government suggested legislation to tighten control over AI-generated misinformation, focusing on openness and responsibility in AI content creation. Public-private partnerships incorporating technology businesses are actively developing detection and watermarking techniques, recognizing the importance of collaborative governance beyond traditional lawmaking.

These selected legal approaches have different emphases: South Korea's emphasis on privacy and punishment, Australia's integrated victim-support system, and Japan's emerging framework that prioritizes AI transparency and industry engagement. Their experiences offer useful insights into developing successful, context-sensitive deepfake law that balances rights protection, technology innovation, and social trust.

INDIA’S ROAD BEYOND: COMPARATIVE PERSPECTIVES-

The worldwide regulatory landscape for deepfake technology provides India with significant comparative insights and lessons that will help it develop an effective legal and legislative response to the particular issues posed by synthetic media. Examining jurisdictions such as the

United States, European Union, China, the United Kingdom, and select Asia-Pacific models highlights similarities and differences that India might use to strengthen its objectionable content regulation and technology safeguards. The effectiveness of targeted, harm-specific regulations tackling non-consensual deepfake pornography, political misinformation, and financial fraud is exemplified by the United States' fragmented yet rapidly growing state legislation. India can utilize this strategy to add dedicated deepfake statutes to existing cyber laws, assuring targeted remedies and enforcement. The US experience also emphasizes the critical role of civil remedies in supplementing criminal punishments and promoting victim empowerment.

The European Union's AI Act provides a sophisticated, harmonized framework that emphasizes transparency, mandatory watermarking, and risk-based governance of AI systems. Adopting these principles can significantly improve India's regulatory architecture, notably in terms of demanding unambiguous disclosures of synthetic content and putting compliance obligations on AI developers and platforms. The EU's emphasis on compliance assessments and human oversight serves as a valuable framework for India to implement accountability systems throughout the AI lifecycle. China's Deep Synthesis Provisions highlight the effectiveness of strict, centralized control centered on mandated labeling, user verification, and content tracking. Although India's democratic setting needs better protection of expression and privacy rights, the Chinese model's emphasis on real-name user registration and technical watermarking can improve India's existing legal framework, particularly in terms of preventing anonymity-facilitated abuses. The UK's Online Safety Act exhibits proactive regulation of non-consensual deepfakes and malevolent synthetic content by combining criminal penalties with robust platform responsibility enforced by a designated regulator. India's future techno-legal crackdown may follow this integrated enforcement model, which combines criminal legislation with regulatory monitoring to ensure timely material removal and victim reparation. Finally, South Korea, Australia, and Japan provide different perspectives on how to balance privacy, victim care, and public education with legal reform, all of which are critical in India's complex digital economy. Public-private partnerships and digital literacy efforts in these countries demonstrate the importance of technology collaboration and public awareness in strengthening legislative measures.

India has the potential to benefit from creating a hybrid legal framework that combines the United States' targeted statutes, the European Union's transparency and risk-driven mandates,

China's labeling and tracking rigor, and the United Kingdom's enforcement regime, all while incorporating substantial privacy safeguards and encouraging collaborative governance. Such a personalized model will better address the numerous societal, economic, and political damages that deepfakes cause, building a resilient digital environment that supports innovation while firmly mitigating misuse.

LEGAL CHALLENGES AND POLICY GAPS-

Deepfake technology raises a complicated set of legal issues and policy gaps that India must solve immediately in order to effectively deter misuse while supporting innovation. The creation and spread of deepfake information is fundamentally global and multinational, posing a significant legal concern. It is difficult for India to bring charges against criminals because many of them operate from other countries, exploiting legal blind spots and a lack of cross-border enforcement cooperation. Additionally, anonymity offered by decentralized platforms and encryption technology makes attribution challenging, which makes jurisdictional claims and law enforcement action more challenging. Evidentiary criteria are still another significant obstacle. Significant forensic expertise is needed to demonstrate that content is truly AI-generated with malevolent intent, something that many investigative agencies do not currently possess. Indian courts are still formulating guidelines for the acceptance of AI-based digital evidence, putting a significant burden on victims to demonstrate manipulation. The lack of mandatory AI content labeling laws in India reduces detection transparency, leaving victims and regulators in the dark about synthetic media proliferation.

The slow pace of cybercrime investigations and court proceedings makes compliance considerably more difficult, allowing harmful deepfakes to spread unchecked for a long period of time. While existing legislation, such as Sections 66D, 67, and 69A of the Indian Information Technology Act, 2000 and other clauses in the Bharatiya Nyaya Sanhita, 2023 offer some legal support, they are not always sufficient to resolve deepfake-specific harms such as digital consent violations or identity fabrication. Constitutional and human rights limits necessitate a difficult balance between preventing abuse and preserving freedom of expression. Article 19 of the Indian Constitution provides free speech, although it is subject to reasonable restrictions based on defamation, public order, and morality. Comprehensive regulation of deepfakes must carefully manage these safeguards to avoid disproportionate censorship or chilling impacts on valid creative, educational, or political speech. Article 21's privacy rights

also require safeguards against the unauthorized use of biometric data and synthetic likenesses, as well as effective data protection and consent procedures. Deepfake technology can spur commercial, educational, and creative innovation, but if it is used carelessly, it can cause social upheaval, financial losses, and a decline in confidence.

A sophisticated framework that promotes transparency, user empowerment, and platform accountability while permitting adaptive technology solutions for detection and resolution should be incorporated into India's legal reforms. In order to create flexible yet effective governance that upholds rights while fostering digital growth, multi-stakeholder engagement involving the government, public society, technology developers, and academia is essential. India confronts significant jurisdictional, evidential, and enforcement obstacles in dealing with deepfake issues, which are exacerbated by constitutional guarantees that necessitate careful regulatory calibration. To overcome these obstacles, a strong legal ecosystem that can strike a balance between advancement and protection in the era of synthetic media must be fostered through international cooperation agreements, AI-specific legal norms, the development of technological capacity, as well as dynamic policy discourse.

CONCLUSION AND RECOMMENDATIONS-

India urgently needs to reform its regulatory framework in order to adequately address the problems caused by deepfakes. The Ministry of Electronics and Information Technology (MeitY) is currently consulting on legislative proposals that call for the creation of dedicated deepfake laws as well as amendments to existing statutes, such as the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021. These revisions offer a broad legal definition of "synthetically generated information," which includes AI-generated audio, video, and images, as well as clear rules requiring deepfake content to be prominently labeled and metadata embedded. Labels should cover at least 10% of visual or audio content to ensure transparency and aid consumers in identifying synthetic media. The amendments also include traceability and accountability requirements, which require intermediaries and content hosts to keep records, verify user declarations regarding synthetic content, and remove illegal deepfakes within 36 hours of notice to avoid harms such as misinformation, defamation, and fraud. Improving the regulatory and institutional frameworks is equally important. Existing authorities, such as CERT-In, the Indian Cyber Crime Coordination Centre (I4C), and Grievance Appellate Committees (GACs), require additional resources, technical competence,

and judicial support to detect, investigate, and punish deepfake violations. Strengthening collaboration between central and state law enforcement, as well as dedicated cyber forensic facilities, can improve enforcement effectiveness. On the technology front, India must incentivize platforms to use cutting-edge detection methods such as AI-powered anomaly detection, watermarking, digital signatures, and real-time flagging systems. Algorithms should be devised to verify authenticity while protecting privacy and reducing false positives. To combat the spread of deepfakes, platforms must share threat intelligence and cooperate transparently with law enforcement. Public policy should promote awareness, digital literacy, and victim support. Education efforts should teach consumers how to spot synthetic content, explain the legal remedies available, and report abuses confidently. Victims of deepfake harassment, blackmail, or defamation will benefit from specialized victim counseling and legal help services.

Finally, given the global character of synthetic media, India's reform program must prioritize international cooperation and multi-stakeholder involvement. Harmonizing legal definitions, sharing best practices, and developing cross-border investigation frameworks can help to overcome jurisdictional barriers. Collaboration with global technology companies, civic society, and academics will build a balanced, multidisciplinary approach to deepfake governance—one that promotes innovation while protecting democracy, privacy, and security. India's reform roadmap calls for an integrated strategy that combines targeted legislation, institutional capacity building, technological innovation, public empowerment, and global partnerships to effectively mitigate the multifaceted risks posed by deepfake technology in the digital age. Significant progress has been made in India's deepfake regulatory path, but there are still important areas that need more reform. With India set to enact strict regulations requiring precise legal definitions, prominent labeling of AI-generated content, and enforced platform accountability under the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2025, key findings underscore the growing prevalence and societal impact of synthetic media. By requiring watermarks on at least 10% of visual or audio content and metadata embedding, these regulations seek to improve user transparency and fight disinformation, fostering public confidence while preserving democratic integrity.

Concluding thoughts emphasize that regulating deepfakes requires a careful balancing act between preventing abuse and promoting digital innovation. The future necessitates flexible,

multi-layered governance that combines strong institutional frameworks, technological detection tools, clear legislation, and broad public awareness. Clear guidelines for AI content authenticity and platform due diligence must be given top priority by policymakers while upholding fundamental liberties like privacy and speech. In order to ensure justice without stifling innovation, the judiciary plays a crucial role in impartially interpreting emergent statutes. Stakeholders, including civil society, tech developers, and digital platforms, must work together proactively in detection, education, and redressal mechanisms. In order to address the cross-border issues inherent in synthetic media, urgent calls to action include enshrining comprehensive, deepfake laws to supplement existing statutes; increasing funding for cyber forensic capabilities and enforcement agencies; requiring interoperable detection and labeling technologies by platforms; encouraging digital literacy campaigns, particularly among vulnerable populations; and fostering international cooperation. India can only effectively combat deepfake dangers and capitalize on AI's revolutionary potential with a concerted, forward-thinking strategy that embraces accountability and openness. In spite of the significant obstacles presented by deepfake technology and the broader AI landscape, this multifaceted approach will enable India to create a resilient, inventive, and rights-respecting digital future.

REFERENCES

1. Rob Cover, Deepfake culture: the emergence of audio-video deception as an object of social anxiety and regulation, 36 JOURNAL OF MEDIA & CULTURAL STUDIES 4 (2022).

2. Andrew Ray, Disinformation, Deepfakes and Democracies: The Need for Legislative Reform, 44 U.N.S.W.L.J. 983 (2021).

3. Julia Hollingsworth, *Indian Women Politicians Face Relentless Trolling Online, Report Says*, *CNN (Jan. 22, 2020),* https://edition.cnn.com/2020/01/22/india/india-women-politicians-trolling-*amnesty-asequals-intl/index.html.*

4. Michael Sumner, "Deepfake Disclosure Laws: Global Approaches 2024", *available at*: https://www.scoredetect.com/

5. William Eritrean, "Deep fakes in the AI act", *available at:* https://schjodt.com/

6. Giulia Interesse, "China to Regulate Deep Synthesis (Deepfake) Technology Starting 2023", *available at:* https://www.china-briefing.com/

7. Online Safety act, 2023, *available at:* https://www.legislation.gov.uk/ukpga/2023/50

8. Meera Srikant, "Bharatiya Laws Against Deepfake Cybercrime Opportunities and Challenges", *available at:* https://www.vifindia.org/article/2025/april/28/Bharatiya-Laws-Against-Deepfake-Cybercrime-Opportunities-and-Challenges

9. Saniya Sayyed, "The Deepfake Dilemma: Legal Challenges and Regulatory Frameworks in the Age of Synthetic Media", *available at:* https://lawfullegal.in/the-deepfake-dilemma-legal-challenges-and-regulatory-frameworks-in-the-age-of-synthetic-media/