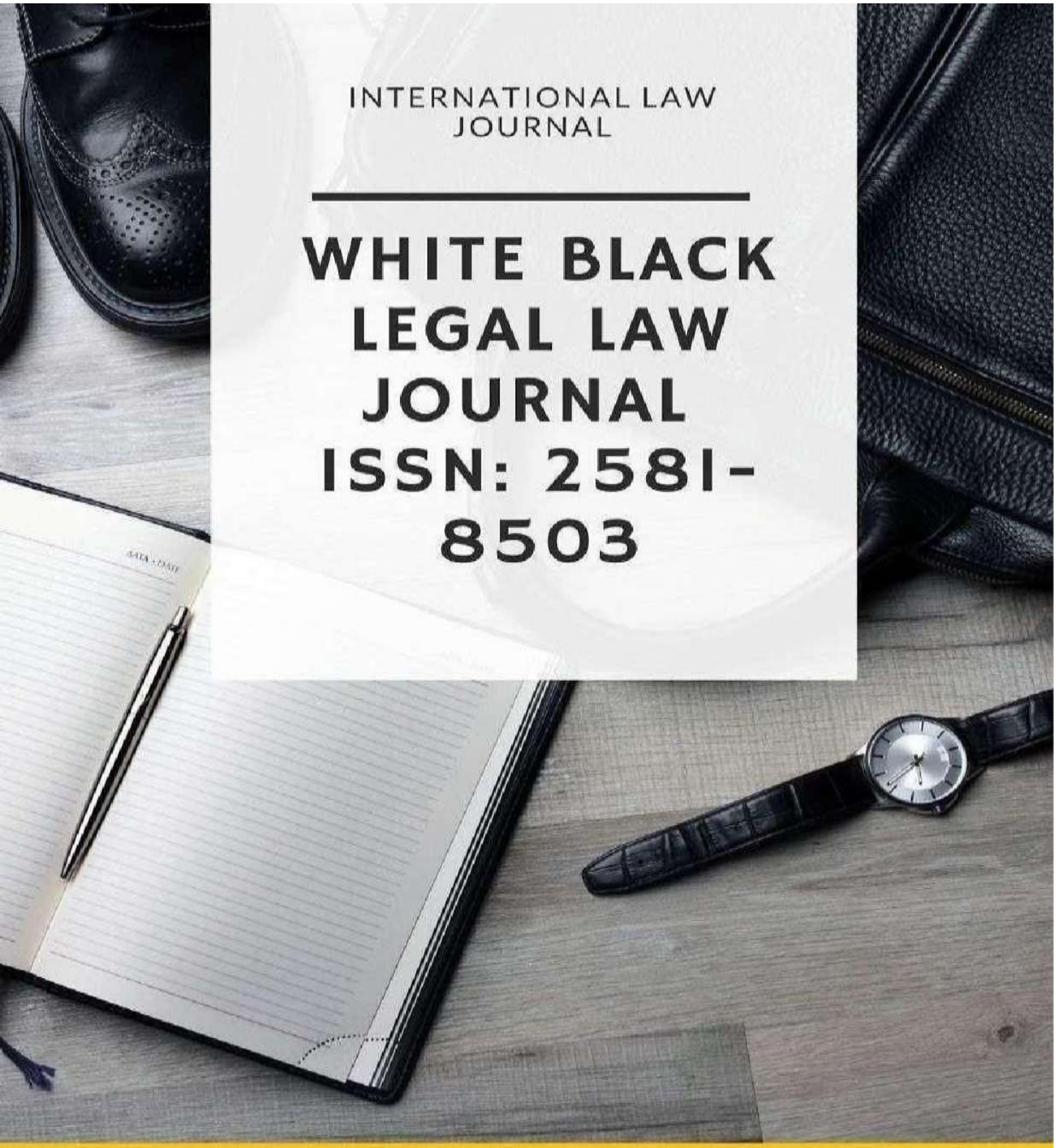

INTERNATIONAL LAW JOURNAL

WHITE BLACK LEGAL LAW JOURNAL

ISSN: 2581-8503

Peer - Reviewed & Refereed Journal

The Law Journal strives to provide a platform for discussion of International as well as National Developments in the Field of Law.

DISCLAIMER

ABOUT WHITE BLACK LEGAL

White Black Legal – The Law Journal is an open-access, peer-reviewed, and refereed legal journal established to provide a scholarly platform for the examination and discussion of contemporary legal issues. The journal is dedicated to encouraging rigorous legal research, critical analysis, and informed academic discourse across diverse fields of law.

The journal invites contributions from law students, researchers, academicians, legal practitioners, and policy scholars. By facilitating engagement between emerging scholars and experienced legal professionals, *White Black Legal* seeks to bridge theoretical legal research with practical, institutional, and societal perspectives.

In a rapidly evolving social, economic, and technological environment, the journal endeavours to examine the changing role of law and its impact on governance, justice systems, and society. *White Black Legal* remains committed to academic integrity, ethical research practices, and the dissemination of accessible legal scholarship to a global readership.

AIM & SCOPE

The aim of *White Black Legal – The Law Journal* is to promote excellence in legal research and to provide a credible academic forum for the analysis, discussion, and advancement of contemporary legal issues. The journal encourages original, analytical, and well-researched contributions that add substantive value to legal scholarship.

The journal publishes scholarly works examining doctrinal, theoretical, empirical, and interdisciplinary perspectives of law. Submissions are welcomed from academicians, legal professionals, researchers, scholars, and students who demonstrate intellectual rigour, analytical clarity, and relevance to current legal and policy developments.

The scope of the journal includes, but is not limited to:

- Constitutional and Administrative Law
- Criminal Law and Criminal Justice
- Corporate, Commercial, and Business Laws
- Intellectual Property and Technology Law
- International Law and Human Rights
- Environmental and Sustainable Development Law
- Cyber Law, Artificial Intelligence, and Emerging Technologies
- Family Law, Labour Law, and Social Justice Studies

The journal accepts original research articles, case comments, legislative and policy analyses, book reviews, and interdisciplinary studies addressing legal issues at national and international levels. All submissions are subject to a rigorous double-blind peer-review process to ensure academic quality, originality, and relevance.

Through its publications, *White Black Legal – The Law Journal* seeks to foster critical legal thinking and contribute to the development of law as an instrument of justice, governance, and social progress, while expressly disclaiming responsibility for the application or misuse of published content.

ISSUES AND CHALLENGES RELATED TO DEEPFAKE TECHNOLOGY IN INDIA: AN EMPIRICAL INVESTIGATION WITH SPECIAL REFERENCE TO NCT OF DELHI

AUTHORED BY - MR. SAGAR MURAL
Faculty of Law, SRM University, Delhi-NCR, Sonepat

1. Introduction

The rapid advent of digital technology in the 21st century has profoundly transformed the human society by altering communication, governance, business, education and social interaction in diverse spheres. The proliferation of internet access, social media and other digital applications throughout the globe have undoubtedly intensified and magnified the scale of exchange of information. Concomitant to this advancement there has also been an alarming rise of new dimensions of cyber threats challenging the rule of law, institutional governance and societal trust in digital platforms. One of the most alarming such advancement being the advent of the deepfake technology.

The term 'deepfake technology' refers to synthesized media, that is to say audio, video, or image content that has been manipulated or generated using an artificial intelligence and deep learning algorithms. Artificial Intelligence-based sophisticated machine learning models create an impersonated digital representation where the person in the synthesized media could be shown saying or performing acts that are false and never did, while in case of voice cloning it involves replicating voice pattern, pitch and intonation of the person so precisely, making deepfakes different from all other previous forms of media manipulation in terms of their scale, sophistication and access.

The continuous misuse of the deepfake technology has thus posed severe threats on a multitude of frontiers namely cyber fraud, identity theft, financial fraud, false impersonation, political propaganda, cyber defamation, reputational damage and nonconsensual sexual media. Individuals ranging from common citizens to politicians and leaders to women are becoming victims of deepfakes. Furthermore, the ability to convincingly depict what is untrue has also

had a massive impact on public trust and confidence towards digital media leading to greater concerns about misinformation, democratic manipulation and the erosion of truth.

In the case of India, the concept of deepfake technology holds extreme significance owing to the exponential increase in its digital population. Internet and smartphone users have increased considerably in the largest market in the world, i.e. India. Coupled with the availability of the internet on nominal prices and participation of the citizens in social networking platforms, this has amplified the pace of digital integration and consequently created fertile grounds for misuse and circulation of false and manipulated content. Varying degrees of digital literacy of the people also enhance their vulnerability towards false and fabricated information, resulting in vulnerability towards financial fraud and other cyber threats.

Although the threat of deepfake technology is alarming, there is no existing separate legal framework that can deal with a synthesized media offence. Various statutes that may be considered to tackle such offenses are the IT Act, 2000, Bharatiya Nyaya Sanhita and Digital Personal Data Protection Act, 2023, all of which only contain some indirect provisions for privacy invasion, cybercrimes, defamation, or cyber impersonation. Such laws were not made to cater to the particular technological aspects, legal challenges and jurisdictional issues arising from deepfake offenses and the resultant legal intervention still seems to be retrospective rather than preventive, creating significant gaps for both the victims and institutions.

Whereas numerous scholarly researches have already been conducted worldwide on technological and regulatory aspect of deepfake technology, research concerning deepfake technology in India still seem to be in a nascent stage, especially, with regard to the empirical analyses in respect of public awareness, institutional readiness, and challenges faced in their enforcement. This indicates the urgency of comprehensive research concerning deepfake technology on legal as well as societal sphere in India.

The aim of the present study therefore, is to critically examine the legal and policy challenges and issues pertaining to the deepfake technology by following a both doctrinal as well as empirical method, particularly on the context of National Capital Territory of Delhi. While scrutinizing the lacunas in existing laws, technological risks, lack of public awareness and limitations in the institutions it seeks to recommend future legislative changes, policy formulation and strategies for effective governance in the digital space against deepfake

technology in India.

2. Literature Review

In the past decade, deepfake has received increasing attention among scholars because of its multifaceted implications for law, cyber security, digital governance, ethics, privacy and democratic stability. Existing literature has established deepfakes as more than just technological innovations, but serious threats to societal trust, institutional legitimacy and legal regimes. Indian and global literature on deepfakes reveals a growing recognition of both their potential and danger across technical, social, legal and political domains.

The earliest body of scholarly writing has concerned an explanation and understanding of the conceptual framework and technical structure of deepfake technology. Mika Westerlund, in his 2019 article *'The Emergence of Deepfake Technology: A Review'* published in Technology Innovation Management Review provides a conceptual overview[1] and one of the earliest academic expositions of deepfakes as a consequence of artificial intelligence and deep learning advances. Westerlund argues that deepfakes are essentially disruptive digital technologies that are poised to alter communications by allowing the easy creation of incredibly realistic synthetically generated media. In terms of its potential, Westerlund posits that whilst it is likely to have beneficial uses in the fields of media, entertainment, education and therapy, misuse by unscrupulous individuals has created severe risks with regards to privacy, trust and security. In a similar vain, Robert Chesney and Danielle Keats Citron's landmark 2019 article *'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security'* published in the California Law Review framing deepfakes as "posing serious threats to privacy, democracy, and national security,"[2] they identify an "information crisis" in the ubiquity of deepfakes in their potential to compromise the integrity of visual and auditory information sources.

The technical literature on deepfakes has predominantly focused on their creation and the development of their detection mechanisms. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio published a paper in 2014 titled *'Generative Adversarial Nets'* which was presented at the Conference on Neural Information Processing Systems (NeurIPS), wherein they introduced a framework for machines to generate highly realistic synthetic information, thus laying the theoretical foundations for modern deepfake systems[3]. Since then, Luisa Verdoliva's article

"Media Forensics and DeepFakes: An Overview" in the IEEE Journal of Selected Topics in Signal Processing[4] highlighted the perpetual arms race between deepfake generation systems and deepfake detection technologies that exist today, stating that: *'deepfake detection techniques are perpetually chasing an ever-moving target'*. Pavel Korshunov and Sbastien Marcel's 2018 article *'DeepFakes: A New Threat to Face Recognition? Assessment and Detection'* presented at the International Conference on Biometrics Engineering and Applications argued that deepfakes pose serious security risks through its capacity to thwart even biometric identification systems such as face recognition[5].

Psychological, reputational, and social harm are increasingly forming a larger part of deepfake scholarship. Luciano Floridi and Massimo Chiriatti in their 2020 article *'GPT-3: Its Nature, Scope, Limits, and Consequences'* in Minds and Machines, discuss that the rapid growth of synthetic media and AI undermines epistemic trust-confidence in the reliability of knowledge and information-which has far reached impacts for the news media, governance, and public debate[6]. Further to the damage of trust and the concept of an information crisis, Chesney and Citron, also discuss the various social damages of deepfakes. These include, importantly, reputational damage through misinformation privacy intrusions and the distribution of non-consensual pornographic deepfake imagery primarily targeted at women. Scholars argue that deepfake harassment against women causes them psychological harm and societal stigma, with devastating social and reputational consequences.

Politically and democratically, deepfakes have been framed as potentially destabilizing forces. In their 2020 article *'Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust'* Cristian Vaccari and Andrew Chadwick explore how manipulated video imagery can be weaponized for the purposes of electoral interference, disinformation campaigns and influencing voter behaviour[7]. Specifically, to India, Gupta and Sinha (2022) in their paper *'Deepfake Technology and the Threat to India's Democratic Ecosystem'* point to India's digital ecosystem as an increasingly vulnerable target for politically motivated deepfakes and how, owing to its vast digital population and immense use of social media, such threats pose a grave danger to the stability of democratic elections and to the public's confidence in government institutions[8].

Legal literature globally has identified significant regulatory gaps with regard to the enforcement of rules and regulations against deepfake usage. Comparative international

analysis shows jurisdictions such as China and the European Union having already begun enacting legislation in relation to deepfakes, including mandated labelling of synthetic content, rules for intermediaries, and standards for the use of AI in general. An example of such laws includes that discussed by Qiang Zhang[9] in *'China's Regulatory Approach Towards Deepfake Technologies'*. In contrast, the legislative framework in India is yet to be adequately addressed with scholars such as Kapoor and Raval and Sharma[10] in their respective works arguing that Indian cyber law remains deficient and that the existing legislation such as the Information Technology Act, 2000 (now replaced by the Bharatiya Nyaya Sanhita, 2023) and the Digital Personal Data Protection Act, 2023 only provide limited remedies rather than adequately addressing the peculiar legal issues arising from the use of deepfakes[11]. This regulatory lacuna makes it difficult to establish liability, provide for protection of victims and to implement procedural and evidentiary mechanisms.

Empirical research carried out in India reveals a distinct lack of awareness amongst the Indian public regarding the technology itself. Whilst Indian social media users may be increasingly encountering deepfakes and doctored media in general, the public is largely not aware of the potential threat and how to identify the use of deepfake technology. Studies carried out by Indian cyber policy organizations and digital rights organizations have indicated that low digital literacy makes Indians highly vulnerable to being manipulated by misinformation through deepfake imagery and to a higher risk of cyber fraud and reputational damage. India's law enforcement bodies, too, have been documented to possess limitations such as the need for increased investment in forensic capacity building, enhancing the technical capabilities and the capacity of law enforcement bodies, and addressing existing procedural difficulties in investigating cyber offenses.

Compared to the already expanding scholarship globally, there is considerably less research conducted within the Indian context. So far, research has treated deepfakes under the more general concept of cybercrime without adequate examination of their unique technological, constitutional, societal and institutional aspects in India. In particular, research remains insufficient when considering aspects such as the public's perception and awareness and how institutional frameworks and legal regimes cope with the use of deepfake technology and there has been very little research carried out to explore these dimensions particularly in urban areas of India like Delhi.

Hence, this research endeavours to bridge this knowledge gap and to offer a critique of deepfake technology in India by adopting an approach of doctrinal legal analysis and combining it with empirical research on the issue. By focusing on the lack of legal infrastructure, vulnerability of the public, and the limitations on the institutional and legal capacity within India, this study seeks to contribute to the nascent body of knowledge required to develop robust and comprehensive governance frameworks for deepfake technology in India.

3. Research Objectives and Methodology

The goal of this research paper is to critique the multi-faceted legal, technical, social and institutional concerns that deepfake technology is poised to throw up in India. Amid the growing threats that deepfake technology poses to privacy, data security, democracy, and public trust, this study aims to narrow the existing gap between technological advancement and legal readiness by examining both doctrinal and empirical aspects of the matter.

The specific objectives of the present study include critically examining the concept and malicious use of deepfake technology, evaluating the adequacy of the present Indian legal and institutional framework in this regard, critically assessing the awareness and susceptibility of the general public to harms emanating from deepfakes, highlighting the real-world challenges faced by law enforcement agencies and proposing practical, viable and legally sound policy and legal reform mechanisms that can govern the use and misuse of synthetic media in India. This framework of objectives aims at providing a holistic view of the impact that deepfake technology is expected to have on an individual, society and the governance structure of the nation.

This objective is achieved by employing a hybrid methodology, namely, a combination of the doctrinal and empirical approaches to research methodology. This approach helps in providing a balance between the purely theoretical concepts of law and its practical application in society.

The doctrinal methodology focuses on critically examining existing legislations, constitutional provisions, precedents and policy initiatives relating to deepfake regulation in India. Relevant legislations like the Information Technology Act, 2000, Bharatiya Nyaya Sanhita, the Digital Personal Data Protection Act, 2023, and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 have been critically assessed and analyzed. Constitutional concerns surrounding privacy, dignity, right to freedom of speech, and the right

to freedom of thought and informational autonomy have been discussed considering the novel problems thrown up by synthetic media[12]. Comparitive reference to various legal and policy frameworks put in place globally, including the European Union, China and the United States, has been attempted to identify best practices.

The empirical methodology relies on primary data gathered by means of a structured questionnaire distributed to 210 social media users across Delhi, the National Capital Territory of Delhi. The purpose of this survey was to collect primary data relating to the public's understanding, susceptibility, awareness, the risks associated with deepfakes, access to legal remedies, confidence in legal and institutional responses, among others. Delhi being chosen as the location for the empirical study helps to focus the research on issues surrounding online media in the urban context in India.

A simple random sampling method was adopted for widespread representation among respondents and the survey was administered through online modes to ensure greater ease of administration and better response rates. Analysis of the data has been done through percentage based interpretation and graphical representations which helped in identifying patterns and trends. Thematic interpretation was used to understand the overall perceptions regarding legal adequacy, institutional trust and policy reforms in the area of deepfake governance.

Secondary sources in the form of books, academic journals, policy papers, government reports, legal databases and institutional documents like NCRB and CERT-In reports have been used to gather secondary data which helped in a deeper theoretical and practical understanding of the challenges in deepfake governance[13].

The combination of these research methods provides the study with both reliability and depth of information, helping it to focus not just on theoretical problems but also on the realistic social context while offering a coherent and evidence-based opinion.

4. Understanding Deepfake Technology and Its Risks

Deepfake is arguably one of the more advanced and concerning developments in the fields of Artificial Intelligence and Digital Media, defining any digital content which is synthetic (audio, video or image) but realistically produced by a system of deep learning models, neural networks

or machine learning algorithms to impersonate another real individual quite convincingly. Unlike previous methods of digital media alteration, deepfakes make use of AI-based generation of artificial content they produce artificial media that are increasingly complex and easier to produce and difficult to identify.

The foundation of deepfake technology lie in the complex computational models, such as the Generative Adversarial Networks (GAN) designed in 2014 by Ian Goodfellow[14] GANs consist of two competing neural networks- one being the 'generator' of the artificial content and other being 'discriminator' to identify the originality of the output. By continuously competing the results gets better and better to produce extremely realistic artificial content that successfully mimic facial movements, voice, posture and speech of the target individual.

Deepfake comes in many varieties and risks, face-swapping deepfake uses manipulation to swap one person's face in the video with another person's face, this kind of deep fake are used to impersonate others in video, non-consensual adult videos and many others. Voice cloning deepfakes copy an individual's vocal characteristic such as their voice tonality, accent and speech to create impersonations and financial scams. Lip-sync deepfakes involve manipulating spoken words, such that an individual appears to say things they did not. Moreover, other advanced types of synthetic media can create completely fictitious individuals without existing human personas the array of potential applications of deepfake technology has made it increasingly both an innovation and an exploitation technology.

Though positive uses for deepfake exist, including for entertainment, film, virtual reality, education, and accessibility, use of the technology in a harmful way has resulted in critical legal, ethical, and societal problems including digital privacy, identity theft, cyber fraud, blackmail, disinformation campaigns for the election process, non-consensual adult material (revenge pornography), reputational harm, and financial fraud and deception. The very nature of a fabricated but realistic human identity creates risks because while the individuals in the media are not real or the event never actually took place they can cause real damage to reputations and social standing.

It can be argued that one of the dangerous issues concerning the usage of deepfake technology is its capability to damage the trust put in digital data. For centuries, an audio or visual representation was a trusted Evidence in law, politics and journalism. In the new paradigm of

synthetic media, this trust in digital media itself is questioned. This phenomenon does not only have an impact on the lives of individuals, but is also poses risk to decision-makers in governance or organizations that rely heavily on accurate information presented through the digital media to function or conduct affairs of the state.

From political stand point, deepfakes has been the key technology for manipulating election process, spreading misinformation and distorting facts to influence the decision of millions. A fabricated political figure saying misleading things could very easily sway the public opinion in real-time on digital platforms. The use of deepfakes for the purpose of harming democratic integrity or political stability is particularly concerning for national security.

Socially, most deeply affected communities with deepfakes technology are vulnerable groups such as women who are typically targeted with non-consensual explicit synthetic content. Such manipulation can take a heavy toll on the victim, causing emotional and psychological trauma, reputational damage, and social ostracism. Professional fields and the public are at high risk of identity manipulation for fraud and reputational attack.

In the Indian context, these deepfakes technology has become much more dangerous due to the country's expanding internet users, burgeoning internet infrastructure, and diverse level of digital literacy. Easily accessible smart phone and affordable internet has enabled easy access to online and social platforms, leading to potential wide and rapid spread of manipulative content at one hand, while many people remain uninformed on how much digital media can be manipulated on other.

Ease of availability of deepfakes making tools has magnified this threat previously a technically well-informed and resourceful individual would require extensive skills to master deepfakes technology, now even the less technical people have the benefit of readily available and user-friendly platforms and software making deepfakes accessible for almost anyone.

Law enforcement and judicial authorities face many challenges trying to keep pace with this development in synthetic media technology. The ability to prove authenticity and veracity of digital media could be a great difficulty and this will give law enforcement authorities more problems in bringing offenders to book for their deeds that use technology. While advanced detection technology is being researched and tested, it will likely lag behind development of

such systems for producing such digital materials.

In essence, deepfake technology is a multidimensional issue which involves technology, law, society, and governance and though technological innovation is advancing at a rapid pace, deepfakes create profound threats to privacy, security, democratic ideals and social cohesion and thus effective strategies involving legislation and governance will be important to counter the potential risks that are becoming more widespread and sophisticated day by day.

5. Legal Challenges and Regulatory Gaps in India

India's existing legal and regulatory framework has been seriously undermined by the advent of deepfake technology. Although India possesses a multitude of cyber laws, penal statutes, and data protection principles, these legislative mechanisms have not been adequately formulated to deal with the multifarious and evolving threats presented by AI-generated synthetic media. As a result, current measures to address the negative consequences of deepfakes remain fragmented, reactive, and inadequate.

Currently, India has no specific law or regulation governing the concept of deepfakes, and only provides for indirect recourse in law. There are no specific statutory provisions that define, classify or regulate the phenomenon of deepfake technology. Existing legislative provisions of the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita and the Digital Personal Data Protection Act, 2023 offer some forms of remedy. However, while certain aspects of deepfakes might be covered within these provisions, such as those concerning fraud, identity theft, defamation, obscenity, impersonation or privacy invasion, they do not fully capture the technological, evidentiary and jurisdictional complexities involved.

Under the Indian Information Technology Act, 2000, Sections 66C and 66D address identity theft and cheating by personation using computer resources respectively and section 67 concerns obscene content in electronic form[15]. While applicable for situations of fraud and impersonation and non-consensual explicit deepfakes, these sections fail to create synthetic media as a distinct legal concept, thereby leading to problems in application and enforcement. Moreover, they were conceived much before the recent surge in AI technologies.

Indian Penal Code or its modern replacement, Bharatiya Nyaya Sanhita, includes criminal

provisions dealing with defamation, fraud, harassment and forgery, but again they do not deal specifically with the manipulation of digital content as presented by deepfakes. Deepfakes are not directly covered under existing definitions and remedies under criminal law as they have a unique identity theft, reputational injury and misinformation component to them, beyond traditional criminal law frameworks.

The Digital Personal Data Protection Act, 2023 does enhance data protection but it would not effectively combat deepfakes as the offence does not always entail direct data theft or misuse and images/videos used could be publicly available. This leaves the domain of privacy protection for synthetic media under this act somewhat inadequate[16].

There are several Constitutional concerns arising from the deployment of deepfakes within India's existing legal framework. Unauthorized alteration or manipulation of someone's image, video or audio, clearly violates the rights to privacy, dignity, reputation and informational autonomy of an individual that has been established by India's Supreme Court in cases like *Justice K.S. Puttaswamy v. Union of India*[17]. Similarly, deepfake defamation issues are interconnected with the right to dignity and personal liberty under Article 21, while the counterpoint here is the tension between the need to regulate it and the Right to Freedom of Speech and Expression under Article 19(1)(a).

One of the most crucial legal challenges with deepfake regulation lies in evidentiary authenticity. Traditionally, audio and video evidence in courts is considered fairly persuasive. However, the increasing sophistication of deepfakes undermines the integrity of digital evidence. This creates major issues in terms of procedurals before courts, law enforcement and lawyers to ascertain authenticity, require expertise, advanced tools and specific protocols which are currently lacking.

Jurisdictional complexities are also quite extensive, as content can be created anonymously, spread rapidly across multiple platforms, hosted outside of Indian borders. This renders cross-border investigation and prosecution highly problematic, and poses significant obstacles given the lack of specific enforcement tools for deepfake crimes at the transnational level.

Institutional preparedness in terms of specialized knowledge, digital forensic infrastructure, tools and uniform procedures for detection, investigation and prosecution remain woefully

inadequate for India's law enforcement agencies in handling the constantly evolving threat of deepfakes. International experience further shows India's lagging regulatory approaches while China has mandate labels for synthetic content, and the EU AI framework promotes transparency and risk regulation[18].

Therefore, India's current legal framework fails to adequately address the concerns and threats posed by deepfakes specifically, because of the absence of specific legislation, piecemeal application of various statutes, constitutional ambiguities and tensions, challenges with evidence, inadequate enforcement tools, institutional incapacitation, and cross-border jurisdictional issues. The way forward necessitates a comprehensive reform of laws, clear legislative definitions of deepfakes, enhanced protections for victims, accountability of intermediaries and sophisticated legal mechanisms capable of handling the latest advancements in synthetic media technology.

6. Empirical Findings

The empirical component of this research involved a structured survey administered to 210 social media users within the National Capital Territory of Delhi. The purpose of this empirical study was to assess public awareness, perception, exposure and susceptibility in relation to deepfake technology while simultaneously examining respondent understanding of legal remedies, institutional readiness and overall societal anxieties. The survey results offer important practical insight into the practical challenges and consequences of deepfake technology in India and complement the doctrinal legal analysis by highlighting existing social and institutional deficits.

The demographics of the survey participants consisted of digitally-active internet users with daily social media consumption, making them ideal respondents for assessing the practical risks associated with synthetic media. Given that users of social media are most likely to be exposed to manipulated digital content in India, their responses provided a direct measure of public awareness and vulnerability within India's evolving digital ecosystem.

Based on the findings, awareness of deepfake technology among the general Indian public was only at a moderate level. While most respondents had heard of deepfake technology or had some familiarity with the concept, they did not have an accurate or precise understanding of

how the technology is used, how sophisticated it could be or what were the actual legal and societal implications. Awareness of the term "deepfake" itself is certainly growing but Digital literacy relating to synthetic media was not developed enough. It was therefore evident that the hypotheses that the awareness level about the concept and legal issues related to deepfake is low and not adequate would hold true.

The results indicate that while the vast majority of respondents had encountered Manipulated or suspicious digital content, the majority of respondents were unsure of their ability to discern authentic digital content from digitally manipulated media. A significant proportion of the respondents lack confidence in their ability to correctly identify deepfakes, indicating a strong vulnerability to misinformation, fraud and reputational damage. It was clear that lower digital literacy and a lack of understanding of synthetic media correlates with increased susceptibility. Privacy was the central most significant finding of the study. The majority of respondents believed that deepfake technology posed a threat to personal privacy, human dignity and reputation. Most of the respondents also acknowledged that freely available pictures, videos and voice records of themselves could be misused by technology, by impersonation, fraud, blackmail or defamation. The findings strongly point towards widespread concerns relating to violations of privacy and identity.

The research results show that a large portion of respondents are unaware of legal remedies and institutional structures in place to assist victims of deepfake misuse. This was demonstrated by a considerable number of respondents being unable to provide details of specific laws, cybercrime complaint mechanisms or relevant institutions to help tackle deepfake-related harm. These findings suggest that even when victims identify digital harm, they may not have adequate awareness regarding the legal protection they can avail themselves of. This reemphasizes problems of legal inaccessibility and institutional unpreparedness.

Regarding the adequacy of the current legal framework in India to curb deepfake related offenses, the majority of the respondents considered the current laws inadequate. Majority respondents thought that the current law does not adequately protect against synthetic media misuse. This finding reflects the doctrinal observations of the fragmentation, indirectness and reactive nature of the Indian legal framework against synthetic media offenses[19].

A very significant empirical finding of the research was the decline in public trust in digital

media. A substantial portion of the respondents indicated that the prevalence of digital manipulation has had a negative effect on the credibility they attach to information encountered online and communication via social media platforms as well as the trustworthiness of digital evidence. The negative impact this had on individuals and on the confidence they had in journalism, democratic processes and information ecosystems as a whole, indicates the growing crisis of digital credibility.

The overwhelming majority of the respondents stated that their expectations from current legislations, platform regulation and technological development are very high. The respondents desire more legislative support, stringent regulation from technology platforms, more advanced technological mechanisms for identification and effective awareness campaign about the growing menace of deepfake.

To conclude, the empirical findings demonstrated that deepfake technology is a serious threat, posing legal, social and institutional risks in India. While awareness regarding this concept and its related legal implications remains only at a moderate level among the masses, they remain at high risk to misinformation, fraud and reputational damage due to their lack of technical awareness, strong concerns about their privacy and a clear lack of confidence in the ability to detect manipulated media. The findings of the study confirmed the lack of legal redressal, institutional preparedness and the dire need for stringent and immediate measures by policy makers to combat the emerging challenges from synthetic media.

7. Suggestions and Recommendations

In light of the rising instances of misuse of deepfake technology and the legal, social and institutional issues elucidated in the doctrinal and empirical analysis it is high time India should take a comprehensive and forward-looking approach in this regard. The first and foremost recommendation is the introduction of legislation which should be solely dedicated to deal with the technology of deepfake. While sections 66E and 67C of the Information Technology Act, 2000, sections of the Bharatiya Nyaya Sanhita and section 7(a) and section 8 of the Digital Personal Data Protection Act, 2023, provide some fragmented remedy for synthetic media related offenses but, lack proper provisions for specific deepfake offenses. Thus, a separate law, explicitly defining deepfakes, punishing wrongful creation, dissemination and amplification and holding creators and distributors accountable, should be introduced. Special

and effective legal remedies should be provided for the victims of deepfakes.

Second, the intermediary and platform accountability needs to be strictly reinforced. Digital content sharing websites, social media platforms and all other intermediaries providing internet access should be legally bound to adopt strict content moderation, detection and fast takedown mechanisms for all maliciously created and circulated deepfakes. Moreover, the law should mandatorily provide for the labelling of AI generated/synthetic content. Regulatory framework of EU and China can be used as model wherein they have started pushing for a 'synthetic media transparency obligation'[20].

Third, India has to significantly increase its law enforcement and digital forensics capacities. Efficient investigation into deepfake offenses would require advanced infrastructure and expert training, thus, dedicated cyber-crime units with latest AI detection capabilities for effective identification of manipulated content are required. There should be increased emphasis on capacity building and constant up gradation of the police forces and related agencies.

Fourth, the level of awareness and digital literacy regarding deepfakes has to be improved. The survey conducted on social media users showed moderate level of awareness with the general tendency to believe in them which makes it a cause of concern[21]. It is important to launch large scale campaigns to make citizens aware about risks, methods to detect them, mechanisms to safeguard privacy and about the remedies available under law. Digital literacy courses must incorporate sections about the risks and appropriate ways of interacting with synthetic media.

Fifth, specialized protection for vulnerable groups of individuals, specifically women and all other victims, should be established. A mechanism that allows for the protection of their identity, has quick complaint filing mechanisms and ensures that adequate psychological support and strong punishment is meted out for sexual exploitation through deepfakes should be implemented.

Sixth, cross border cooperation is inevitable as the sources of deepfakes can be outside the jurisdiction and spread through digital channels which are globally accessible. Strengthening international legal collaborations and cyber-crime agreements is necessary to deal with international ramifications of deepfake technology[22].

Seventh, regular policy research and institutional monitoring are indispensable. The rapid evolution of artificial intelligence necessitates a proactive regulatory approach thus, constant engagement and coordination between government agencies, academia and research institutions should be encouraged for assessing emerging risks and policy recommendations.

8. Conclusion

It can be seen that deepfake technology is among the most important concerns arising out of interaction between AI, electronic communication and governance, on the modern scene. While such advancements in technology open vast possibilities, yet at the same time, deepfake technology raises some of the worst risks that individuals face – from those of privacy, defamation, cybersecurity to democracy and trustworthiness of digital media. The risks in the Indian context become graver on the account of large digital ingress in the country which shows varying level of digital literacy coupled with absence of an exhaustive regulatory framework specifically dealing with synthetic media.

This research established that, indeed, the challenges posed by deepfake technology in India are multifaceted and cover legal vacuum, institutional limitations, citizen's vulnerability and social damage. We established that present laws are not enough, fragmented and often retrospective and are inadequate to deal with the special techno-evidentiary and legal problems posed by synthetic media. We find from our findings that awareness among the public, among other respondents of our research is moderate. There is also found a relatively weak legal understanding, coupled with immense privacy concerns, decreasing trust on digital media, coupled with strong desire of people to reform laws.

We have seen that it is not only individuals who bear the brunt of the harm caused due to deepfakes-their potential to disturb the democratic discourse with misinformation and weaken the trustworthiness of digital communication and future of legal regime based on audio visual evidence-will take some time to be properly comprehended by human beings. With the continuous advancements in synthetic media and greater accessibility it poses a major threat that the lag between development of technology and legal readiness will keep growing wider unless serious policy changes are made.

India should adopt a more holistic approach in this domain by combining reformative

legislative measures, technological countermeasures, institutional robustness, platform liability, enhanced public awareness, international cooperation in order to tackle emerging risks from deepfakes efficiently, and to ensure the benefit derived from innovation[23]. It can thus be seen that deepfake technology not only impacts individual rights, social balance and the nature of democratic discourse but is a pressing social and legal challenge requiring robust policy intervention. The level of future safety and credibility of Indian digital ecosystem will depend on it.

Bibliography

Books, Articles and Journals

Chesney Robert and Danielle Keats Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107 *California Law Review* 1753.

Floridi Luciano and Massimo Chiriatti, 'GPT-3: Its Nature, Scope, Limits, and Consequences' (2020) 30 *Minds and Machines* 681.

Goodfellow Ian J, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio, 'Generative Adversarial Nets' in *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)* (2014).

Gupta and Sinha, 'Deepfake Technology and the Threat to India's Democratic Ecosystem' (2022) *Journal of Information, Law and Technology* 52.

Kapoor and Raval, 'Regulating Deepfakes in India: Legal Challenges and Policy Responses' (2021) *Indian Journal of Law and Technology* 85.

Korshunov Pavel and Sébastien Marcel, 'DeepFakes: A New Threat to Face Recognition? Assessment and Detection' in *Proceedings of the International Conference on Biometrics Engineering and Applications* (2018).

Sharma, 'Deepfake Technology and Indian Cyber Law: A Critical Analysis' (2023) 6(3) *International Journal of Law Management & Humanities* 233.

Vaccari Cristian and Andrew Chadwick, 'Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust' (2020) 6(1) *Social Media + Society* 1.

Verdoliva Luisa, 'Media Forensics and DeepFakes: An Overview' (2020) 14(5) *IEEE Journal of Selected Topics in Signal Processing* 910.

Westerlund Mika, 'The Emergence of Deepfake Technology: A Review' (2019) 9(11) *Technology Innovation Management Review* 39.

Zhang Qiang, 'China's Regulatory Approach Towards Deepfake Technologies' (2021) *Computer Law & Security Review* 105531.

Legislations

Bharatiya Nyaya Sanhita, 2023.

Digital Personal Data Protection Act, 2023. Information Technology Act, 2000.

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.

Cases

Justice K.S. Puttaswamy (Retd.) v Union of India (2017) 10 SCC 1.

Reports and Government Documents

European Parliament, *Artificial Intelligence Act* (2024), available at: https://artificialintelligenceact.eu/

Indian Computer Emergency Response Team (CERT-In), *Annual Report 2023*, available at: https://www.cert-in.org.in

National Crime Records Bureau, *Crime in India Report 2023* (Ministry of Home Affairs, Government of India, 2023), available at: https://ncrb.gov.in

United Nations Office on Drugs and Crime (UNODC), *Comprehensive Study on Cybercrime* (2024), available at: https://www.unodc.org/

[1] Mika Westerlund, 'The Emergence of Deepfake Technology: A Review' (2019) 9(11) *Technology Innovation Management Review* 39, 40–42, available at: https://timreview.ca/article/1282

[2] Robert Chesney and Danielle Keats Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107 *California Law Review* 1753, 1758–1765, available at: https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security/

[3] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio, 'Generative Adversarial Nets' in *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)* (2014) 2672, 2672–2680, available at: https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[4] Luisa Verdoliva, 'Media Forensics and DeepFakes: An Overview' (2020) 14(5) *IEEE Journal of Selected Topics in Signal Processing* 910, 912–918, available at: https://ieeexplore.ieee.org/document/9078746

[5] Pavel Korshunov and Sébastien Marcel, 'DeepFakes: A New Threat to Face Recognition? Assessment and Detection' in *Proceedings of the International Conference on Biometrics Engineering and Applications* (2018) 1, pg. 1–6, available at: https://arxiv.org/abs/1812.08685

[6] Luciano Floridi and Massimo Chiriatti, 'GPT-3: Its Nature, Scope, Limits, and Consequences' (2020) 30 *Minds and Machines* 681, pg. 690–693, available at: https://link.springer.com/article/10.1007/s11023-020-09548-1

[7] Cristian Vaccari and Andrew Chadwick, 'Deepfakes and Disinformation: Exploring the Impact of Synthetic

Political Video on Deception, Uncertainty, and Trust' (2020) 6(1) *Social Media + Society* 1, pg. 4–10, available at: https://journals.sagepub.com/doi/full/10.1177/2056305120903408

[8] Gupta and Sinha, 'Deepfake Technology and the Threat to India's Democratic Ecosystem' (2022) *Journal of Information, Law and Technology* 52, pg. 56–61.

[9] Qiang Zhang, 'China's Regulatory Approach Towards Deepfake Technologies' (2021) *Computer Law & Security Review* 105531, 105533–105537, available at: https://www.sciencedirect.com/science/article/pii/S0267364921000507

[10] Sharma, 'Deepfake Technology and Indian Cyber Law: A Critical Analysis' (2023) 6(3) *International Journal of Law Management & Humanities* 233, 240–245, available at: https://www.ijlmh.com/

[11] Kapoor and Raval, 'Regulating Deepfakes in India: Legal Challenges and Policy Responses' (2021) *Indian Journal of Law and Technology* 85, 90–97.

[12] *Justice K.S. Puttaswamy (Retd.) v Union of India* (2017) 10 SCC 1, 497–500.

[13] National Crime Records Bureau, *Crime in India Report 2023* (Ministry of Home Affairs, Government of India, 2023), available at: https://ncrb.gov.in ; Indian Computer Emergency Response Team (CERT-In), *Annual Report 2023*, available at: https://www.cert-in.org.in

[14] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio, 'Generative Adversarial Nets' in *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)* (2014) 2672–2680, available at: https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[15] Information Technology Act, 2000, ss 66C, 66D and 67.

[16] Digital Personal Data Protection Act, 2023, ss 7(a) and 8

[17] *Justice K.S. Puttaswamy (Retd.) v Union of India* (2017) 10 SCC 1, 497–500.

[18] Qiang Zhang, 'China's Regulatory Approach Towards Deepfake Technologies' (2021) *Computer Law & Security Review* 105531, 105533–105537, available at: https://www.sciencedirect.com/science/article/pii/S0267364921000507

[19] Kapoor and Raval, 'Regulating Deepfakes in India: Legal Challenges and Policy Responses' (2021) *Indian Journal of Law and Technology* 85, 90–97; Sharma, 'Deepfake Technology and Indian Cyber Law: A Critical Analysis' (2023) 6(3) *International Journal of Law Management & Humanities* 233, 240–245.

[20] Qiang Zhang, 'China's Regulatory Approach Towards Deepfake Technologies' (2021) *Computer Law & Security Review* 105531, 105533–105537, available at: https://www.sciencedirect.com/science/article/pii/S0267364921000507 ; European Parliament, *Artificial Intelligence Act* (2024), available at: https://artificialintelligenceact.eu/

[21] Cristian Vaccari and Andrew Chadwick, 'Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust' (2020) 6(1) *Social Media + Society* 1, 4–10, available at: https://journals.sagepub.com/doi/full/10.1177/2056305120903408

[22] United Nations Office on Drugs and Crime (UNODC), *Comprehensive Study on Cybercrime* (2024), available at: https://www.unodc.org/

[23] European Parliament, *Artificial Intelligence Act* (2024), available at: https://artificialintelligenceact.eu/ (last visited 13 May 2026); Qiang Zhang, 'China's Regulatory Approach Towards Deepfake Technologies' (2021) *Computer Law & Security Review* 105531, pg. 105533–105537.