CONSTITUTIONAL PROTECTION AGAINST DEEPFAKE SEXUAL CONTENT IN INDIA: ANALYSING PRIVACY, DIGNITY AND CRIMINAL LIABILITY

P. River, BA LLB (H), Amity University, Madhya Pradesh

ABSTRACT

The proliferation of deepfake technology - the use of artificial intelligence and machine learning algorithms to create hyper-realistic but entirely fabricated digital content - has given rise to one of the most alarming and most rapidly growing categories of digital sexual violence in contemporary society. The non-consensual creation, publication, and dissemination of deepfake sexual content causes devastating psychological, reputational, and social harm to its victims, who are disproportionately women and marginalised persons, and raises profound questions of criminal liability, constitutional protection, and regulatory adequacy that existing Indian legal frameworks are ill-equipped to address. This paper undertakes a comprehensive doctrinal and analytical examination of the criminal liability and constitutional dimensions of deepfake sexual content in India, assessing the adequacy of the existing legal framework - including the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Digital Personal Data Protection Act, 2023, and the constitutional protections of Articles 19 and 21 of the Constitution of India - and identifying the specific legislative, institutional, and constitutional reforms needed to provide effective protection against AI-generated sexual exploitation. The paper advances the central argument that the absence of specific legislation governing deepfake sexual content in India constitutes a serious and urgent constitutional deficit that leaves victims without adequate legal remedies and perpetrators without appropriate criminal accountability, and proposes a comprehensive framework of legislative, institutional, and judicial reforms to address this deficit.

Keywords: Deepfakes, Artificial Intelligence, Digital Sexual Violence, Criminal Liability, Right to Privacy, Article 21, Bharatiya Nyaya Sanhita 2023, Information Technology Act, Intermediary Liability, Constitutional Protection, Digital Dignity.

I. INTRODUCTION

The intersection of artificial intelligence and human sexuality has produced one of the most disturbing and most legally challenging phenomena of the digital age - the creation and dissemination of deepfake sexual content. Deepfake technology, which employs deep learning algorithms and generative adversarial networks to superimpose a person's face or likeness upon the body of another person in sexually explicit material, has made it possible for any person with a smartphone and a free application to create convincingly realistic fabricated sexual imagery of any individual whose photographs are available online. The consequences for victims - who are overwhelmingly women, with studies indicating that over ninety percent of all deepfake pornography online targets women without their consent - are catastrophic and often irreversible, encompassing severe psychological trauma, professional destruction, reputational devastation, and in some cases suicidal ideation.[1]

India, with the world's largest population of internet users and one of the highest rates of smartphone penetration globally, presents a particularly acute context for the harms of deepfake sexual content. The combination of widespread digital literacy, the pervasive use of social media platforms, deeply entrenched patriarchal social norms, and the relative anonymity afforded by digital communication creates an environment in which the non-consensual creation and dissemination of deepfake sexual imagery can cause exceptional and sustained harm to its victims. Yet the Indian legal framework - as most recently updated by the Bharatiya Nyaya Sanhita, 2023, the Bharatiya Sakshya Adhiniyam, 2023, and the Digital Personal Data Protection Act, 2023 - contains no specific provisions addressing deepfake sexual content and no statutory definition of AI-generated synthetic media, leaving prosecutors, courts, and victims to navigate a complex and inadequate patchwork of provisions originally designed for entirely different categories of harm.[2]

The present paper is structured as follows. Section II examines the technology of deepfakes and the specific harms generated by non-consensual deepfake sexual content. Section III analyses the existing Indian legal framework for the regulation of deepfake sexual content, including the provisions of the IT Act, the BNS 2023, and the constitutional framework. Section IV examines the criminal liability dimensions of deepfake sexual content, identifying the specific offences applicable and the gaps in criminal accountability. Section V examines the constitutional dimensions of deepfake sexual content, with particular reference to the right to dignity and privacy under Article 21 and the right to equality under Article 14. Section VI

examines the intermediary liability framework and content moderation challenges. Section VII undertakes a comparative analysis of legal responses to deepfake sexual content in selected international jurisdictions. Section VIII advances specific reform recommendations. Section IX concludes.

II. DEEPFAKE TECHNOLOGY AND THE NATURE OF THE HARM

2.1 The Technology of Deepfakes

Deepfake technology derives its name from the combination of deep learning - a subset of machine learning based upon artificial neural networks - and the word fake, reflecting the technology's capacity to create realistic fabrications indistinguishable from genuine media. The technical foundation of deepfake generation is the Generative Adversarial Network (GAN), a machine learning architecture in which two neural networks - a generator and a discriminator - compete against each other in a process that iteratively improves the quality and realism of the generated output. The generator creates synthetic media by learning the visual and auditory characteristics of the target individual from training data - typically photographs or videos available online - and combining those characteristics with the base content to create a composite that appears authentic.[3]

The technical barriers to the creation of convincing deepfakes have fallen dramatically in recent years. Applications that were once the exclusive domain of well-resourced technology companies and research laboratories are now available as free or low-cost consumer applications downloadable from major app stores. FaceApp, DeepFaceLab, Reface, and numerous other applications allow users to create sophisticated face-swapping and image manipulation content with minimal technical expertise, using smartphones as the only required hardware. The democratisation of deepfake creation technology has created an unprecedented scale of potential harm - any person whose photographs appear online, which in the age of social media means virtually every adult and many children in India, is a potential target for non-consensual deepfake sexual content creation.[4]

2.2 The Specific Harms of Deepfake Sexual Content

The harm caused by non-consensual deepfake sexual content is distinctive in character and exceptional in severity in ways that distinguish it from other categories of digital harm and that demand a specifically tailored legal response. Unlike conventional defamatory content,

deepfake sexual content creates a false but visually compelling impression that the victim actually participated in the depicted sexual activity, making denial and refutation extremely difficult. Unlike stolen intimate images - the category of harm addressed by revenge pornography legislation - deepfake sexual content does not require access to genuine intimate imagery of the victim, making it accessible to any person with access to the victim's publicly available photographs.[5]

The psychological impact of deepfake sexual content upon its victims has been documented extensively in the academic literature. Studies from jurisdictions where victims have been surveyed report rates of post-traumatic stress disorder, severe depression, anxiety disorders, and suicidal ideation that are comparable to those associated with physical sexual assault. The permanence of digital content - the fact that once shared, deepfake sexual material may be impossible to completely remove from the internet - means that the psychological harm to victims is not time-limited but may persist for years or decades. The professional consequences of deepfake sexual content can include job loss, loss of professional reputation, and enforced withdrawal from public life that have particularly severe implications for women in public-facing professional roles including politicians, journalists, academics, and public advocates.[6]

The gendered character of deepfake sexual harm is a critically important dimension that the legal framework must acknowledge and address. Research consistently shows that the overwhelming majority of deepfake pornography targets women, and that the creation and dissemination of such content is frequently motivated by misogyny, the desire to silence women in public and professional discourse, revenge following relationship breakdown, and the exercise of control and intimidation. The use of deepfake sexual content as a tool of gender-based violence and digital abuse places it squarely within the international human rights framework for the elimination of violence against women, and the failure of the Indian legal framework to specifically address this form of gender-based digital violence is a matter of international human rights concern as well as a domestic legal deficiency.

III. THE EXISTING INDIAN LEGAL FRAMEWORK

3.1 The Information Technology Act, 2000

The Information Technology Act, 2000 provides the primary statutory framework for cybercrime regulation in India and contains several provisions that are potentially applicable to deepfake sexual content, though none of them was specifically designed with deepfakes or

AI-generated synthetic media in mind. The most directly applicable provisions are Sections 66C, 66D, 67, 67A, and 67B of the IT Act, each of which addresses a distinct aspect of the harm caused by deepfake sexual content but none of which provides comprehensive coverage of the full range of harms involved.

Section 66C criminalises identity theft - the fraudulent or dishonest use of another person's electronic signature, password, or any other unique identification feature. The application of Section 66C to deepfake sexual content is arguable insofar as the creation of a deepfake involves the fraudulent use of a person's facial identity - a biometric characteristic that constitutes a unique identification feature - but the provision was designed primarily for financial identity theft and its application to the specific harm of deepfake sexual content is uncertain and contested.[7]

Section 66D criminalises cheating by personation through a communication device - using a computer resource or communication device to cheat by impersonating another person. The creation of deepfake sexual content could potentially fall within the ambit of Section 66D insofar as it involves the impersonation of the victim in explicitly sexual contexts, but the provision requires proof of an intent to cheat - a mens rea requirement that may not be satisfied in all cases of deepfake sexual content creation, particularly where the primary motive is harassment or degradation rather than financial gain.

Sections 67 and 67A criminalise the publication or transmission of obscene material and sexually explicit material respectively in electronic form. These provisions are potentially applicable to the distribution of deepfake sexual content - which is by definition sexually explicit - but their focus on publication and transmission means that they do not address the harm of creation itself and do not capture cases where deepfake sexual content is created but not yet disseminated. The definition of sexually explicit material in Section 67A is also not specifically tailored to the characteristics of AI-generated synthetic media and may raise definitional difficulties in prosecutions involving deepfake content.[8]

Section 67B addresses the creation, publication, or transmission of child sexual abuse material in electronic form and provides the most comprehensive and clearly applicable IT Act provision for cases involving deepfake sexual content targeting children. The provision explicitly covers material depicting children in obscene or sexually explicit manner and is not limited to genuine footage of actual abuse, making it potentially applicable to deepfake sexual

content depicting minors. However, the protection of Section 67B is expressly limited to minors and provides no corresponding protection for adult victims of deepfake sexual exploitation.

3.2 The Bharatiya Nyaya Sanhita, 2023

The Bharatiya Nyaya Sanhita, 2023 - which replaced the Indian Penal Code, 1860 on 1st July 2024 - contains several provisions that may be applicable to deepfake sexual content, primarily in the areas of defamation, sexual harassment, stalking, voyeurism, and criminal intimidation. However, like the IT Act, the BNS 2023 contains no specific provision addressing deepfake technology or AI-generated synthetic media, leaving victims to rely upon general provisions that were designed for different categories of harm and that may not provide adequate protection in the specific context of deepfake sexual exploitation.

Section 356 of the BNS 2023 addresses criminal defamation and may be applicable where deepfake sexual content is created and disseminated with the intent to harm the victim's reputation. The constitutional validity of criminal defamation was upheld by the Supreme Court in Subramanian Swamy v. Union of India,[9] where the Court affirmed that reputation is an integral component of the right to life under Article 21. The application of Section 356 to deepfake sexual content requires proof of the intent to harm reputation - a requirement that may be satisfied in many cases of deepfake sexual exploitation but that does not capture all the harmful dimensions of such content, including cases where the primary motivation is sexual gratification, harassment, or control rather than reputational damage.

Section 77 of the BNS 2023 addresses voyeurism and specifically includes the capture, publication, or transmission of images of a private nature through electronic means. The application of Section 77 to deepfake sexual content is problematic because the provision appears to contemplate the capture or recording of actual private conduct, rather than the AI-generated fabrication of private conduct that never actually occurred. Whether a deepfake video depicting a person engaged in sexual activity constitutes an image of a private nature within the meaning of Section 77, when the depicted conduct is entirely fabricated, is a question of statutory interpretation that has not yet been addressed by the courts.[10]

Section 78 of the BNS 2023 addresses stalking and specifically includes monitoring a person's use of the internet, email, or electronic communication. This provision may be applicable

where deepfake sexual content is used as a tool of persistent harassment within a pattern of stalking behaviour, but does not provide a standalone basis for criminalising the creation or dissemination of deepfake sexual content outside the stalking context.

The BNS 2023 provisions on sexual harassment under Section 75 and criminal intimidation under Section 351 may also be applicable in specific cases of deepfake sexual exploitation, where the content is created and used as a means of intimidation, coercion, or harassment of the victim. However, these provisions, like the others examined, address specific aspects of the harm caused by deepfake sexual content without providing comprehensive coverage of the full range of criminal conduct involved.

3.3 The Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act, 2023 establishes a framework for the protection of personal data in India that has direct relevance to deepfake sexual content, as the creation of deepfake imagery involves the unauthorised processing of the victim's biometric data - specifically their facial characteristics - for purposes that the victim has not consented to and that are fundamentally incompatible with their autonomy and dignity. The DPDP Act requires that personal data be processed only for lawful purposes and with the consent of the data principal, and its application to the processing of biometric data for the purpose of creating deepfake sexual content would appear to be straightforward in principle.[11]

However, the DPDP Act's provisions are primarily directed at data fiduciaries - entities that process personal data for defined purposes within a regulated commercial framework - and are less well-adapted to the context of individual wrongdoers who process another person's biometric data for the purpose of creating harmful content. The enforcement mechanisms of the DPDP Act - which centre on data protection authority oversight and administrative penalties - are also less well-suited to the criminal justice context in which deepfake sexual exploitation cases primarily arise, and the interaction between the DPDP Act's data protection framework and the criminal liability framework of the BNS 2023 and the IT Act has not yet been clarified by judicial interpretation or legislative guidance.

IV. CRIMINAL LIABILITY FOR DEEPFAKE SEXUAL CONTENT

4.1 The Actus Reus - Applicable Conduct

The criminal liability framework applicable to deepfake sexual content in India must address

three distinct categories of criminal conduct - the creation of deepfake sexual content, its distribution and dissemination, and its use as an instrument of coercion, intimidation, or harassment. Each category raises distinct criminal law questions and is addressed with varying degrees of adequacy by the existing legislative framework.

The creation of deepfake sexual content - the act of generating artificial sexual imagery incorporating the likeness of an individual without their consent - is the foundational harm from which all subsequent harms derive. Under the existing Indian legal framework, the act of creation alone - without dissemination - is arguably not addressed by any specific criminal provision. Section 67 of the IT Act criminalises publication and transmission but not mere creation; the voyeurism and stalking provisions of the BNS 2023 contemplate recording or monitoring of actual conduct rather than the creation of fictional imagery. This gap in the criminal framework - the absence of a specific offence of creating non-consensual deepfake sexual content - is one of the most significant deficiencies in the Indian legal response to this form of harm.[12]

The distribution and dissemination of deepfake sexual content - the sharing of created material through digital platforms, messaging applications, or other electronic means - is more clearly addressed by the existing framework through the publication and transmission offences of Sections 67 and 67A of the IT Act. However, the specific challenges of digital dissemination - including the speed at which content spreads, the difficulty of containing dissemination once it has commenced, and the international character of many dissemination networks - create practical enforcement challenges that the legal framework does not adequately address.

The use of deepfake sexual content as an instrument of coercion or blackmail - threatening to create or disseminate such content unless the victim complies with the perpetrator's demands - is addressed by the criminal intimidation provisions of the BNS 2023 and may also constitute extortion under Section 308 of the BNS 2023. These provisions provide a basis for criminal prosecution in cases of sextortion using deepfake material, but their application requires proof of the specific intent to coerce or intimidate that may not be present in all harmful cases.

4.2 The Mens Rea Problem

The determination of the appropriate mens rea for deepfake sexual content offences is a particularly challenging aspect of the criminal liability framework. Different categories of

perpetrators may be motivated by different mental states - sexual gratification, revenge, misogynistic harassment, financial gain through blackmail, or a generalised desire to harm - and the applicable provisions of the IT Act and the BNS 2023 each impose their own specific mens rea requirements that may not correspond to the full range of culpable mental states involved in deepfake sexual exploitation.[13]

The present study argues that a specific deepfake sexual content offence should be drafted as a strict liability offence in relation to the act of creation and dissemination - imposing criminal liability upon any person who knowingly creates or disseminates non-consensual deepfake sexual content without requiring proof of a specific intent to harm - subject to a defence of reasonable belief in the victim's consent. This approach, modelled upon the consent-based framework of modern sexual offence legislation, would provide comprehensive coverage of the full range of culpable conduct while providing an appropriate defence for cases involving genuine consent.

4.3 Evidentiary Challenges

The prosecution of deepfake sexual content offences presents specific and significant evidentiary challenges that arise from the technical character of the evidence and the forensic capabilities required to establish the key elements of the offence. The most fundamental evidentiary challenge is the authentication of deepfake content as artificially generated - establishing that the sexual content depicts a fabrication rather than actual conduct of the victim - which requires the application of sophisticated AI-based forensic detection tools that are not currently available in most Indian forensic laboratories.[14]

The identification of the perpetrator of deepfake sexual content is a further significant evidentiary challenge. The creation and dissemination of deepfake sexual content typically occurs through anonymous or pseudonymous online channels, and the attribution of content to a specific individual may require the forensic analysis of IP addresses, device identifiers, network traffic records, and platform metadata that is subject to the same cross-border jurisdictional challenges examined in the context of cybercrime investigation generally. The admissibility of the resulting digital forensic evidence under the BSA 2023 certification framework is governed by the general principles examined in the present study's main dissertation, and the specific application of those principles to the distinctive evidentiary challenges of deepfake investigation is an area requiring specific judicial guidance.

V. CONSTITUTIONAL DIMENSIONS

5.1 The Right to Privacy and Dignity under Article 21

The constitutional foundation for the protection of individuals against deepfake sexual content is most directly provided by Article 21 of the Constitution of India, which guarantees the right to life and personal liberty and has been interpreted by the Supreme Court to encompass a broad range of substantive rights including the right to dignity, the right to privacy, and the right to informational autonomy. The Supreme Court's landmark decision in Justice K.S. Puttaswamy v. Union of India[15] established that the right to privacy is a fundamental right under Article 21 and that it encompasses the protection of personal identity, informational autonomy, and bodily integrity against unlawful interference by both state and non-state actors.

The application of the Puttaswamy framework to deepfake sexual content is direct and compelling. The creation of non-consensual deepfake sexual content involves the unauthorised manipulation of the victim's facial identity - a core aspect of personal identity - for the purpose of placing them in a degrading and humiliating sexual context that they have neither consented to nor had any opportunity to prevent. This manipulation directly violates the victim's informational autonomy - their right to control how their identity and image are used and represented in the public domain - and constitutes a profound invasion of the personal dignity that is a central component of the right to life under Article 21.[16]

The constitutional dimensions of deepfake sexual harm extend beyond the privacy framework to encompass the right to reputation - which the Supreme Court has recognised as an element of the right to life under Article 21 in Subramanian Swamy v. Union of India[17] - and the right to psychological integrity - which is increasingly recognised in the constitutional jurisprudence as a component of the right to life that encompasses protection against the deliberate infliction of severe psychological harm. The devastating psychological consequences of deepfake sexual victimisation - including post-traumatic stress disorder, severe depression, and suicidal ideation - constitute harms to the constitutional right to life and personal liberty that the state has a constitutional obligation to prevent and remedy through appropriate legislative measures.

5.2 The Right to Equality and Non-Discrimination under Article 14

The gendered character of deepfake sexual harm - the fact that the overwhelming majority of victims are women and that the creation and dissemination of such content is frequently

motivated by misogyny and the desire to silence and control women - raises important questions under Article 14 of the Constitution of India, which guarantees the right to equality before the law and equal protection of the laws. A legal framework that fails to provide adequate protection against a form of violence that is disproportionately directed at women may constitute a denial of the equal protection of the laws to women in violation of Article 14.[18]

The Supreme Court's progressive development of the equality jurisprudence in the context of gender-based violence - including the recognition of the constitutional dimensions of sexual harassment in Vishaka v. State of Rajasthan and the articulation of a positive state obligation to protect women from violence - provides a constitutional basis for arguing that the state's failure to enact specific legislation addressing deepfake sexual violence constitutes a violation of its Article 14 obligations to provide women with equal protection against gender-based harm.

5.3 The State's Positive Constitutional Obligation

The constitutional framework established by the Puttaswamy decision and the broader jurisprudence on fundamental rights imposes not merely a negative obligation upon the state to refrain from violating the right to privacy and dignity but a positive obligation to protect individuals against violations of those rights by third parties, including private individuals and corporations. This positive constitutional obligation requires the state to enact legislation that effectively protects the right to digital dignity and privacy against the harms of deepfake sexual exploitation, and the current legislative vacuum in this area represents a failure of the state to discharge its positive constitutional obligations under Article 21.

The present study argues that the positive constitutional obligation of the state to protect individuals against deepfake sexual exploitation encompasses the obligation to enact specific legislation criminalising the non-consensual creation and dissemination of deepfake sexual content, to establish institutional mechanisms for the investigation and prosecution of such offences, and to provide effective remedies - including compensation and the compelled removal of harmful content - for victims of deepfake sexual exploitation.[19]

VI. INTERMEDIARY LIABILITY AND CONTENT MODERATION

6.1 The Section 79 Safe Harbour and Its Limitations

The intermediary liability framework of Section 79 of the Information Technology Act, 2000 provides the primary legal framework governing the obligations of digital platforms in relation

to harmful content including deepfake sexual material. Section 79 provides that an intermediary shall not be liable for third-party content if it does not initiate the transmission, does not select the recipients of the transmission, and does not select or modify the information contained in the transmission. This safe harbour protection is conditional upon the intermediary complying with prescribed due diligence requirements and taking down harmful content upon receiving actual knowledge of its presence.[20]

The application of the Section 79 safe harbour to deepfake sexual content raises several specific challenges. The due diligence requirements prescribed by the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 require intermediaries to publish clear rules prohibiting harmful content, to establish grievance redressal mechanisms, and to remove content upon receiving a complaint. However, the 72-hour removal timeline prescribed by the 2021 Rules for certain categories of harmful content may be insufficient to prevent the rapid viral dissemination of deepfake sexual content, which can reach millions of viewers within hours of being uploaded.

The technical detection and prevention of deepfake sexual content by platforms presents a specific challenge that the existing intermediary liability framework does not adequately address. The detection of deepfake content - as distinct from genuine sexually explicit material - requires the application of sophisticated AI-based detection tools that many platforms, particularly smaller platforms operating in India, do not currently deploy. The absence of a specific statutory obligation upon intermediaries to deploy deepfake detection technology is a significant gap in the content moderation framework that allows harmful deepfake content to persist on platforms for extended periods despite being technically detectable.

6.2 Platform Accountability and the Proactive Obligation

The present study argues that the existing reactive model of intermediary liability - under which platforms are required to remove harmful content only after receiving actual notice - is inadequate for the specific harms of deepfake sexual content and should be supplemented by a proactive obligation upon large digital platforms to deploy technically effective deepfake detection systems and to prevent the upload of non-consensual synthetic sexual imagery before it is disseminated to users.[21]

This proactive obligation, modelled upon the obligations imposed upon very large online

platforms under the European Union's Digital Services Act, would require major social media platforms, messaging applications, and video sharing services operating in India to implement technically effective content detection systems for deepfake sexual material, to establish rapid victim notification and content removal procedures, and to report regularly to the designated regulatory authority on the volume and nature of deepfake sexual content detected and removed from their platforms.

VII. COMPARATIVE LEGAL RESPONSES

7.1 United Kingdom

The United Kingdom has enacted the most comprehensive legislative framework for non-consensual deepfake sexual content currently in force in any common law jurisdiction. The Online Safety Act, 2023 specifically criminalises the sharing of intimate images without consent - including deepfake intimate images - and the Criminal Justice Bill, introduced in 2023, goes further by creating a specific offence of creating deepfake intimate images without consent, regardless of whether the created material is ever shared. This dual approach - criminalising both creation and sharing as separate offences - provides a comprehensive framework of criminal accountability that addresses the full lifecycle of deepfake sexual harm from creation through dissemination.[22]

The United Kingdom's approach is particularly noteworthy for its recognition that the act of creation - regardless of subsequent sharing - is itself a harm that warrants criminal sanction. This approach reflects a sophisticated understanding of the nature of deepfake harm that goes beyond the visible harm of dissemination to recognise the violation of dignity and autonomy inherent in the creation of non-consensual sexual imagery, and provides a model for the development of comparable legislation in India.

7.2 United States

The United States has adopted a primarily state-level legislative approach to non-consensual deepfake sexual content, with numerous states - including California, Texas, Virginia, Georgia, and New York - having enacted legislation specifically addressing deepfake intimate imagery. California's AB 602 and AB 730 provide civil remedies for victims of deepfake sexual content and criminalise the distribution of deepfake pornography, while Virginia's Code Section

18.2-386.2 specifically criminalises the creation and distribution of deepfake sexually explicit material without consent.[23]

At the federal level, the DEEPFAKES Accountability Act proposes to require disclosure of AI-generated content and to establish criminal penalties for the non-consensual creation and distribution of deepfake sexual content. While the federal legislation has not yet been enacted, its proposal reflects growing legislative recognition at the national level of the need for comprehensive federal regulation of deepfake sexual harm.

7.3 European Union

The European Union's approach to deepfake sexual content operates through multiple overlapping regulatory frameworks - the Artificial Intelligence Act, 2024, which imposes transparency and watermarking obligations on AI systems capable of generating synthetic media, the Digital Services Act, 2022, which imposes content moderation obligations on very large online platforms, and the forthcoming Directive on Combating Violence against Women, which specifically addresses technology-facilitated sexual violence including non-consensual synthetic intimate imagery. The EU's multi-layered regulatory approach - combining AI regulation, platform regulation, and criminal law - provides a model for the comprehensive governance of deepfake sexual harm that is particularly instructive for the development of comparable frameworks in India.[24]

7.4 China

China has enacted the most technically specific and most immediately enforceable regulatory framework for deepfake content currently in force anywhere in the world. The Provisions on the Administration of Deep Synthesis Internet Information Services, 2022, require the mandatory watermarking of all AI-generated content, including deepfakes, before publication online. This watermarking requirement creates a technical trail that enables the identification of AI-generated content and the tracing of its origin, providing both a deterrent to the creation of harmful deepfake content and a forensic tool for the investigation of deepfake offences. The mandatory watermarking approach provides a technically grounded and practically effective complement to criminal law sanctions that India should consider adopting as a component of its regulatory framework.

VIII. REFORM RECOMMENDATIONS

On the basis of the doctrinal and comparative analysis undertaken in the preceding sections, the present paper advances the following specific recommendations for the reform of the Indian legal framework governing deepfake sexual content.

First, the Parliament of India should enact a dedicated Digital Sexual Violence Prevention Act that specifically addresses the creation, dissemination, and use of non-consensual deepfake sexual content. This legislation should provide clear statutory definitions of deepfake technology, synthetic sexual imagery, and non-consensual intimate imagery, create specific criminal offences for the creation and dissemination of non-consensual deepfake sexual content with appropriate sentencing provisions, and establish consent as the central organising principle of the offence framework.[25]

Second, the dedicated legislation should establish a civil remedy framework for victims of deepfake sexual exploitation that provides for injunctive relief - including emergency orders for the removal of harmful content - damages for psychological, reputational, and professional harm, and a victim compensation fund for cases where the perpetrator cannot be identified or is unable to pay damages.

Third, the Information Technology Act, 2000 should be amended to impose a specific statutory obligation upon major digital intermediaries to deploy technically effective deepfake detection systems, to establish rapid victim notification and content removal procedures with a maximum removal timeline of 24 hours from notification, and to report regularly to the Ministry of Electronics and Information Technology on deepfake sexual content detection and removal activities.

Fourth, the government should mandate the watermarking of AI-generated content, modelled upon the Chinese Provisions on the Administration of Deep Synthesis Internet Information Services, 2022, requiring all AI tools and applications capable of generating synthetic media to embed a technically robust and publicly verifiable watermark in all generated content before it is disseminated.[26]

Fifth, the Ministry of Home Affairs should establish dedicated deepfake investigation units within state Cyber Crime Investigation Cells, equipped with AI-based deepfake detection tools,

forensic analysis capabilities for synthetic media, and specialised training in the investigation and prosecution of deepfake offences. These units should be supported by a national technical reference laboratory within the Central Forensic Science Laboratory that maintains current knowledge of deepfake detection methodologies and provides technical assistance to state investigation units.

Sixth, the Supreme Court of India should issue a constitutional declaration that the non-consensual creation and dissemination of deepfake sexual content constitutes a violation of the fundamental right to privacy and dignity under Article 21 and that the state has a positive constitutional obligation to provide effective legal remedies against such violations. Such a declaration would provide the constitutional foundation for the legislative reforms proposed above and would signal the judiciary's commitment to protecting digital dignity in the age of artificial intelligence.

Seventh, a comprehensive public education programme should be developed and implemented - through schools, digital platforms, and mass media - to raise awareness of the harms of deepfake sexual content, the legal obligations of platform users, and the remedies available to victims.[27]

IX. CONCLUSION

Deepfake sexual content represents one of the most serious and most rapidly growing threats to the dignity, privacy, and equality of individuals in contemporary Indian society. The convergence of increasingly accessible AI technology, widespread digital connectivity, and deep-seated misogynistic attitudes creates the conditions for an epidemic of AI-generated sexual exploitation that the existing Indian legal framework - characterised by general provisions designed for pre-AI categories of harm and the complete absence of specific deepfake legislation - is fundamentally inadequate to address.

The constitutional framework of India - with its recognition of the right to dignity and privacy under Article 21, the positive state obligation to protect individuals against violations of fundamental rights, and the guarantee of equal protection under Article 14 - provides not merely the normative foundation for legislative reform but the constitutional imperative for it. The state's continuing failure to enact specific legislation addressing deepfake sexual exploitation is not merely a legislative oversight but a constitutional deficit that leaves the most

vulnerable members of Indian society without adequate protection against one of the most harmful and most pervasive forms of technology-enabled abuse.

The comparative analysis undertaken in this paper has demonstrated that the challenge of deepfake sexual content is not unique to India and that several comparable jurisdictions - including the United Kingdom, the United States, the European Union, and China - have developed legislative and regulatory responses that provide instructive models for reform. The reform recommendations advanced in Section VIII of this paper draw upon these comparative models while adapting them to the specific constitutional framework, institutional realities, and socio-cultural context of India, and constitute a comprehensive and practically grounded agenda for the development of the legal framework that India's digital citizens urgently require.

The protection of dignity in the digital age is not a luxury or a secondary concern - it is a constitutional imperative. The development of a comprehensive legal framework for the governance of deepfake sexual content is therefore not merely a technical legislative task but a statement of the values that Indian society chooses to uphold in the age of artificial intelligence. It is to the making of that statement - through legislation, judicial decision, and institutional reform - that the present paper is dedicated.

BIBLIOGRAPHY

Books

1. Citron, Danielle Keats: *Hate Crimes in Cyberspace*, Harvard University Press, Cambridge, 2014.

2. Duggal, Pavan: *Cyberlaw: The Indian Perspective*, Saakshar Law Publications, New Delhi, 2nd edn., 2002.

3. Kamath, Nandan: *Law Relating to Computers, Internet and E-Commerce*, Universal Law Publishing, New Delhi, 5th edn., 2020.

4. Murray, Andrew: *Information Technology Law: The Law and Society*, Oxford University Press, Oxford, 2019.

5. Nappinai, N.S.: *Cyber Crimes and the Law*, OakBridge Publishing, New Delhi, 2010.

6. Rai, Uday Raj: *Fundamental Rights and their Enforcement*, Eastern Book Company, Lucknow, 3rd edn., 2016.

7. Sharma, Vakul: *Information Technology Law and Practice*, Universal Law Publishing, New Delhi, 5th edn., 2022.

8. Singh, Justice Yatindra: *Cyber Laws*, Universal Law Publishing, New Delhi, 2019.

9. Viswanathan, Aparna: *Cyber Law: Indian and International Perspectives*, LexisNexis, New Delhi, 2021.

10. Zweigert, Konrad and Kötz, Hein: *An Introduction to Comparative Law*, Oxford University Press, Oxford, 3rd edn., 1998.

Statutes

1. The Constitution of India, 1950.

2. The Information Technology Act, 2000 (Act 21 of 2000).

3. The Bharatiya Nyaya Sanhita, 2023 (Act 45 of 2023).

4. The Bharatiya Nagarik Suraksha Sanhita, 2023 (Act 46 of 2023).

5. The Bharatiya Sakshya Adhiniyam, 2023 (Act 47 of 2023).

6. The Digital Personal Data Protection Act, 2023 (Act 22 of 2023).

7. The Protection of Children from Sexual Offences Act, 2012 (Act 32 of 2012).

8. The Indecent Representation of Women (Prohibition) Act, 1986 (Act 60 of 1986).

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

10. The Online Safety Act, 2023 (United Kingdom).

11. Regulation (EU) 2022/2065 of the European Parliament and of the Council (Digital Services Act, 2022).

12. Regulation (EU) 2024/1689 of the European Parliament and of the Council (Artificial Intelligence Act, 2024).

13. People's Republic of China, Provisions on the Administration of Deep Synthesis Internet Information Services, 2022.

Cases

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

2. Selvi v. State of Karnataka, (2010) 7 SCC 263.

3. Shreya Singhal v. Union of India, AIR 2015 SC 1523.

4. Subramanian Swamy v. Union of India, (2016) 7 SCC 221.

5. R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.

6. Vishaka v. State of Rajasthan, AIR 1997 SC 3011.

7. Kharak Singh v. State of Uttar Pradesh, AIR 1963 SC 1295.

8. Anvar P.V. v. P.K. Basheer, AIR 2015 SC 180.

9. Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 SCC 1.

Journal Articles

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

2. Citron, Danielle Keats: "Sexual Privacy" 128 *Yale Law Journal* 1870 (2019).

3. Citron, Danielle Keats and Franks, Mary Anne: "Criminalizing Revenge Porn" 49 *Wake Forest Law Review* 345 (2014).

4. Franks, Mary Anne: "Sexual Harassment 2.0" 71 *Maryland Law Review* 655 (2012).

5. Lahiri, Karan: "Deepfakes, Free Speech and Platform Liability in India" 14 *Indian Journal of Law and Technology* 92 (2022).

6. Das Acevedo, Deepa: "Digital Harassment and Privacy Rights in India" *Indian Journal of Law and Technology* 89 (2021).

7. Groh, Matthew: "Deepfake Detection and the Role of Artificial Intelligence" 12 *Journal of Cyber Policy* 44 (2021).

8. Ajder, Henry: "The State of Deepfakes: Landscape, Threats and Impact" 7 *Deeptrace Report* 12 (2019).

9. Baxi, Upendra: "Human Dignity in Constitutional Governance" 44 *Journal of Indian Law Institute* 321 (2002).

10. Duggal, Pavan: "Cybercrime and Digital Evidence in India" 3 *Supreme Court Journal* 45 (2018).

Reports

1. NITI Aayog: *National Strategy for Artificial Intelligence*, Government of India, New Delhi, 2018.

2. Law Commission of India: *Report No. 267 on Hate Speech*, Ministry of Law and Justice, New Delhi, 2017.

3. National Crime Records Bureau: *Crime in India Report 2022*, Ministry of Home Affairs, Government of India, New Delhi, 2023.

4. UNESCO: *Recommendation on the Ethics of Artificial Intelligence*, Paris, 2021, available at: https://www.unesco.org (last visited on 1st March 2026).

5. Ministry of Electronics and Information Technology: *Principles for Responsible AI*, Government of India, New Delhi, 2021.

Websites

1. https://www.meity.gov.in

2. https://www.ncrb.gov.in

3. https://www.internetfreedom.in

4. https://www.eff.org

5. https://www.sci.gov.in

6. https://www.europarl.europa.eu