IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI-Generated Deepfakes and Cyber Violence Against Women in India: Legal Challenges and Policy Response

Bhawya Prajapati, Ph.D. Scholar, School of Law, Apex University | Dr Surbhi Dadhich, Associate Professor, School of Law, Apex University, Jaipur

Abstract

The emergence of generative Artificial Intelligence has transformed the digital ecosystem by enabling the creation of realistic synthetic images, audio recordings, and videos with minimal human intervention. Although these technological advancements have contributed to innovation across various sectors, they have simultaneously facilitated new forms of cybercrime. Among the most alarming developments is the misuse of deepfake technology to target women through the creation and dissemination of fabricated intimate content, identity impersonation, cyber harassment, online blackmail, and reputational attacks. Such practices threaten women's digital safety and raise serious concerns regarding privacy, dignity, autonomy, and access to justice.

The increasing sophistication and accessibility of deepfake tools present significant challenges for legal systems, particularly in countries such as India where technology is evolving more rapidly than regulatory mechanisms. This paper examines the emerging phenomenon of AI-generated deepfakes as a form of technology-enabled gender-based violence and evaluates the effectiveness of the existing Indian legal framework in responding to these offences. The study critically analyses relevant statutory provisions, judicial developments, and intermediary obligations while identifying practical barriers to investigation, attribution, evidence collection, and prosecution.

Employing a doctrinal research methodology supplemented by contemporary case studies and comparative legal perspectives, the paper argues that existing cyber laws provide only fragmented protection against AI-driven harms. It emphasises the necessity of developing a dedicated legal framework capable of addressing the unique challenges posed by generative AI technologies. The paper further recommends stronger regulatory oversight, enhanced platform responsibility, specialised digital forensic capabilities, improved victim-support mechanisms, and comprehensive public awareness initiatives. It concludes that a balanced legal and policy approach is essential to safeguard women's rights, strengthen cyber resilience, and ensure that

advances in Artificial Intelligence are aligned with constitutional values and the principles of justice.

Keywords: Artificial Intelligence (AI); Generative Artificial Intelligence; Deepfake Technology; Cyber Violence Against Women; Cybercrime; Digital Privacy; Women's Digital Rights; Technology-Facilitated Gender-Based Violence; Information Technology Act, 2000; Bharatiya Nyaya Sanhita, 2023.

1. Introduction

1.1 Background of the Study

Artificial Intelligence (AI) has emerged as one of the most influential technological innovations of the twenty-first century, transforming the manner in which information is created, processed, and disseminated. Among its recent advancements, generative AI has enabled the development of deepfake technology, which uses sophisticated machine learning algorithms to produce highly realistic yet fabricated images, audio recordings, and videos. While deepfakes have legitimate applications in education, entertainment, healthcare, journalism, and digital media, their misuse has become an emerging challenge for cybersecurity, privacy, and the administration of justice.

In India, the widespread availability of AI-powered tools and social media platforms has facilitated the rapid creation and dissemination of manipulated digital content. Women have become one of the most vulnerable groups affected by this technology, with deepfakes increasingly being used for non-consensual intimate imagery, identity impersonation, cyberstalking, online harassment, extortion, and character assassination. Such offences not only violate constitutional rights to privacy, dignity, and equality but also have long-lasting psychological, social, and economic consequences. The increasing incidence of AI-enabled cyber violence necessitates an examination of the adequacy of India's legal and regulatory framework in addressing these evolving threats.

1.2 Statement of the Problem

The rapid evolution of generative AI has significantly outpaced the development of legal and regulatory mechanisms designed to combat technology-enabled crimes. Although existing legislation addresses certain cyber offences, it does not specifically regulate AI-generated deepfakes or comprehensively address the unique challenges associated with synthetic media. Difficulties relating to the identification of offenders, authentication of digital evidence, intermediary liability, jurisdictional issues, and victim protection continue to impede effective enforcement. Consequently, there is an urgent need to evaluate whether the current legal framework sufficiently protects women from AI-enabled cyber violence and to identify reforms capable of addressing emerging technological risks.

1.3 Research Gap

Existing scholarship on Artificial Intelligence and cyber law primarily focuses on the technological aspects of deepfakes, general cybercrime, or the ethical implications of AI. Comparatively fewer studies examine AI-generated deepfakes as a distinct form of gender-based cyber violence within the Indian legal context. Furthermore, limited research critically analyses the effectiveness of recently enacted legislation, including the Bharatiya Nyaya Sanhita, 2023, and the Digital Personal Data Protection Act, 2023, in addressing offences involving AI-generated content. This study seeks to bridge this gap by providing a comprehensive

legal analysis of AI-generated deepfakes, their impact on women's digital rights, and the policy reforms necessary for strengthening cyber governance in India.

1.4 Objectives of the Study

The objectives of this research are:

1. To examine the concept and evolution of AI-generated deepfake technology.
2. To analyse the various forms of cyber violence against women facilitated through deepfakes.
3. To evaluate the effectiveness of the existing Indian legal framework in addressing AI-generated deepfake offences.
4. To identify the legal and practical challenges encountered in the investigation and prosecution of deepfake-related crimes.
5. To recommend legal and policy measures for preventing the misuse of AI-generated deepfakes and strengthening the protection of women's digital rights.

1.5 Research Questions

The study seeks to address the following research questions:

1. What are AI-generated deepfakes, and how are they being used to facilitate cyber violence against women?
2. Does the existing Indian legal framework adequately address offences arising from AI-generated deepfakes?
3. What challenges do law enforcement agencies encounter in investigating and prosecuting deepfake-related cybercrimes?
4. What legal and policy reforms are necessary to regulate AI-generated deepfakes while ensuring technological innovation and the protection of fundamental rights?

1.6 Scope and Significance of the Study

The study focuses on the legal implications of AI-generated deepfakes as an emerging form of cyber violence against women in India. It examines the applicability of existing cyber laws, criminal laws, data protection legislation, intermediary liability provisions, and judicial developments relating to AI-enabled offences. The research also incorporates comparative international perspectives to identify best practices that may strengthen India's regulatory framework.

The significance of this study lies in its contribution to the growing discourse on Artificial Intelligence, cyber law, and women's digital rights. By identifying legislative gaps and proposing practical policy reforms, the research seeks to support policymakers, legal practitioners, researchers, technology developers, and law enforcement agencies in developing a balanced legal framework that promotes technological innovation while safeguarding individual rights and ensuring effective access to justice.

2. Literature Review

Artificial Intelligence (AI) has emerged as a transformative technology with far-reaching implications for governance, business, healthcare, and the legal system. The integration of AI into legal processes has enhanced efficiency in legal research, document analysis, and judicial administration. However, the rapid development of generative AI has also introduced complex legal and ethical challenges, particularly concerning the creation and misuse of deepfake technology (Russell & Norvig, 2021).

Deepfakes are synthetic media generated using deep learning algorithms capable of producing highly realistic images, videos, and audio recordings. While these technologies have legitimate applications in education,

entertainment, and digital communication, they have increasingly been exploited for malicious purposes, including misinformation, identity theft, financial fraud, and cyber violence. Chesney and Citron (2019) argue that deepfakes pose significant risks to privacy, democratic institutions, and public trust because of their ability to manipulate digital content with remarkable realism.

Women have emerged as one of the primary targets of AI-enabled cyber abuse. Technology-facilitated gender-based violence includes non-consensual intimate image creation, cyberstalking, online harassment, impersonation, and image-based sexual abuse. According to Citron (2014), online abuse disproportionately affects women by undermining their privacy, dignity, and professional opportunities. Similarly, McGlynn et al. (2017) emphasise that image-based sexual abuse extends beyond "revenge pornography" and includes any non-consensual creation or distribution of intimate content, including AI-generated deepfakes.

Within the Indian context, existing legal scholarship acknowledges that the Information Technology Act, 2000 provides a legal framework for addressing several cyber offences. However, the Act predates the emergence of generative AI and therefore lacks provisions specifically regulating deepfake technology. Recent legislative developments, including the Bharatiya Nyaya Sanhita, 2023 and the Digital Personal Data Protection Act, 2023, represent important steps towards strengthening cyber governance, yet scholars continue to debate their adequacy in addressing AI-generated harms (Ministry of Electronics and Information Technology [MeitY], 2023).

Recent reports by international organisations further indicate that generative AI has increased the accessibility of sophisticated cybercrime techniques by lowering the technical expertise required to create manipulated content. This development presents significant challenges for law enforcement agencies in identifying perpetrators, authenticating digital evidence, and ensuring effective prosecution (Europol, 2024).

Although previous studies have extensively discussed Artificial Intelligence, cybercrime, privacy, and digital governance, limited research specifically examines AI-generated deepfakes as a form of cyber violence against women within the Indian legal framework. This study seeks to address this gap by critically analysing existing laws, identifying regulatory shortcomings, and proposing legal and policy reforms to strengthen the protection of women's digital rights in the era of generative AI.

3. Research Methodology

This study adopts a doctrinal (qualitative) research methodology to examine the legal and policy challenges posed by AI-generated deepfakes as a form of cyber violence against women in India. Doctrinal research is appropriate because it focuses on analysing legal principles, statutory provisions, judicial decisions, and policy documents to evaluate the adequacy of the existing legal framework governing emerging technologies (Jain, 2016).

The research is primarily based on secondary sources of data. These include national legislation such as the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Digital Personal Data Protection Act, 2023, and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. Judicial decisions delivered by the Supreme Court of India and various High Courts relating to privacy, intermediary liability, cybercrime, and digital rights have also been examined to understand the evolving legal position.

In addition to statutory and judicial sources, the study reviews academic literature, including peer-reviewed journal articles, books, conference proceedings, government reports, publications of international organisations, and policy papers discussing Artificial Intelligence, deepfake

technology, cybercrime, digital privacy, and technology-facilitated gender-based violence. Relevant reports published by organisations such as Europol, UNESCO, and the World Economic Forum have also been consulted to provide comparative and contemporary perspectives on the misuse of generative AI.

The research further employs a comparative legal approach by briefly examining regulatory developments in selected jurisdictions, including the European Union, the United States, and the United Kingdom. This comparative analysis assists in identifying best practices that may inform future legal reforms in India.

The collected information is analysed using a descriptive and analytical approach. The study critically evaluates the strengths and limitations of the existing legal framework, identifies regulatory gaps, and assesses whether current laws adequately address offences involving AI-generated deepfakes. Based on this analysis, the paper proposes policy recommendations aimed at strengthening legal safeguards, improving enforcement mechanisms, enhancing platform accountability, and protecting the digital rights of women in the age of Artificial Intelligence.

Limitations of the Study

The study is confined to a doctrinal analysis of the legal issues surrounding AI-generated deepfakes and cyber violence against women in India. It does not involve empirical data collection through surveys or interviews. Given the rapid evolution of Artificial Intelligence technologies and legal responses, future legislative developments and judicial interpretations may influence the applicability of the findings presented in this paper.

3. Research Methodology

The present study adopts a doctrinal research methodology to examine the legal and policy challenges associated with AI-generated deepfakes as a form of cyber violence against women in India. Since the research primarily focuses on analysing legal principles, statutory provisions, judicial pronouncements, and policy frameworks, the doctrinal method is considered the most appropriate approach for achieving the objectives of the study (Jain, 2016).

The study relies exclusively on secondary sources of data. These include statutes such as the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Digital Personal Data Protection Act, 2023, and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. Judicial decisions, Law Commission reports, government publications, parliamentary debates, peer-reviewed journal articles, books, conference proceedings, and reports published by international organisations have also been examined to provide a comprehensive understanding of the subject.

To place the Indian legal framework in a broader context, the study incorporates a comparative analysis of selected international approaches towards regulating AI-generated deepfakes and technology-facilitated gender-based violence. This comparative perspective helps identify global best practices that may guide future legal reforms in India.

The collected data has been analysed using a descriptive, analytical, and comparative approach. The research critically evaluates the adequacy of existing legal provisions governing AI-generated deepfakes, identifies regulatory and enforcement gaps, and proposes legal and policy recommendations for strengthening the protection of women's digital rights while encouraging the responsible development and use of Artificial Intelligence.

Limitations of the Study

The scope of this research is confined to a doctrinal analysis of the legal and policy dimensions of AI-generated deepfakes and cyber violence against women in India. The study does

not include empirical methods such as surveys or interviews. Furthermore, considering the rapid pace of technological innovation and legal developments in Artificial Intelligence, future legislative and judicial changes may influence the findings and recommendations of this research.

4. Understanding AI-Generated Deepfakes: Concept, Evolution, and Technological Framework

One of the main achievements of Artificial Intelligence (AI) in the digital age is deepfake technology, enabling the creation of highly real yet fake images, audio recordings, and videos. The term deepfake stands for "deep learning + fake" because this technology relies on machine learning and artificial intelligence to generate synthetic media (Chesney & Citron, 2019). AI uses artificial neural networks to accurately reproduce human facial expressions, voices, and body language. In particular, deepfakes are mainly produced with the help of Generative Adversarial Networks (GAN) invented by Goodfellow et al. (2014). This technology involves two neural networks: generator, which produces synthetic content, and discriminator, which distinguishes between the real and fabricated materials. As a result of their repeated training, the generator develops its ability to produce increasingly realistic material, confusing both people and automatic systems (Chesney & Citron, 2019)..

Deepfakes are primarily created using Generative Adversarial Networks (GANs), introduced by Goodfellow et al. (2014). A GAN consists of two neural networks: a generator, which creates synthetic content, and a discriminator, which evaluates its authenticity. Through repeated training, the generator produces increasingly realistic outputs capable of deceiving both human observers and automated detection systems. More recent advancements in diffusion models and other generative AI techniques have further enhanced the quality, accessibility, and speed of synthetic media creation, allowing users with limited technical expertise to generate convincing deepfakes.

The application of deepfake technology is not inherently unlawful. In fact, it has contributed significantly to innovation across multiple sectors. In the entertainment industry, deepfakes are used for visual effects, dubbing, and digital restoration of historical footage. Educational institutions utilise synthetic media to create interactive learning environments, while healthcare researchers employ AI-generated simulations for medical training and patient education. Businesses have also adopted AI-generated avatars and multilingual voice synthesis to improve customer engagement and digital communication (UNESCO, 2023).

Despite these legitimate applications, the rapid proliferation of deepfake technology has created new opportunities for cybercriminals. The widespread availability of open-source AI models and user-friendly software has significantly reduced the technical barriers to producing manipulated content. Consequently, deepfakes are increasingly being exploited for identity fraud, financial scams, political misinformation, corporate espionage, and online abuse. Among these emerging threats, the misuse of deepfake technology to facilitate cyber violence against women has become a matter of serious legal and social concern.

Women are disproportionately targeted through the creation of non-consensual intimate images and videos, impersonation on social media platforms, online harassment, cyberstalking, and digital blackmail. Such misuse extends beyond reputational harm by violating an individual's privacy, dignity, autonomy, and psychological well-being. The viral nature of digital platforms enables manipulated content to spread rapidly, often causing irreversible damage before victims can seek legal remedies. Moreover, the anonymity afforded by cyberspace and the sophisticated nature of AI-generated content create substantial challenges for law enforcement

agencies in identifying offenders and preserving digital evidence.

The growing misuse of AI-generated deepfakes demonstrates that technological innovation must be accompanied by effective legal regulation and ethical governance. While AI continues to offer immense benefits to society, its responsible development requires robust legal safeguards that prevent malicious use without discouraging innovation. Understanding the technological foundations of deepfakes is therefore essential for evaluating the adequacy of existing legal frameworks and developing effective policy responses to protect individuals, particularly women, from AI-enabled cybercrime.

5. AI-Generated Deepfakes as a Form of Cyber Violence Against Women in India

The rapid advancement of generative Artificial Intelligence has transformed the nature of cybercrime by enabling the creation of realistic synthetic media that can be produced and disseminated with minimal technical expertise. Among the most concerning consequences of this technological development is the increasing use of AI-generated deepfakes to perpetrate cyber violence against women. Unlike conventional forms of cybercrime, deepfake technology enables offenders to fabricate highly convincing images, audio recordings, and videos that falsely depict individuals engaging in actions or speech that never occurred. The realistic nature of such content significantly increases its potential to deceive viewers, damage reputations, and cause lasting psychological and social harm (Chesney & Citron, 2019).

Cyber violence against women encompasses a broad range of technology-facilitated offences, including online harassment, cyberstalking, identity theft, image-based sexual abuse, doxxing, impersonation, threats, and the non-consensual dissemination of intimate content. Artificial Intelligence has amplified these offences by making it possible to generate fabricated intimate images and videos using publicly available photographs collected from social media platforms. Victims frequently discover manipulated content circulating online without their knowledge or consent, making removal difficult and often resulting in irreversible reputational damage (Citron, 2014).

One of the most prevalent forms of AI-enabled abuse is the creation of non-consensual intimate deepfakes, in which a person's face is digitally superimposed onto explicit images or videos. Such content is frequently distributed through encrypted messaging services, social media platforms, and illicit websites with the intention of humiliating, intimidating, blackmailing, or exploiting victims. The consequences extend beyond online spaces, affecting victims' mental health, educational opportunities, employment prospects, personal relationships, and overall sense of security. Studies have consistently shown that women are disproportionately targeted by image-based sexual abuse, reflecting existing patterns of gender inequality in digital environments (McGlynn et al., 2017).

In addition to image-based abuse, AI-generated voice cloning has emerged as another significant concern. Advances in speech synthesis now allow perpetrators to imitate an individual's voice with remarkable accuracy using only a short audio sample. Such technology has been used to facilitate fraud, impersonation, extortion, and the dissemination of false information. When combined with deepfake videos, voice cloning further complicates the verification of digital evidence and increases the likelihood of deception.

India has witnessed a noticeable increase in public awareness regarding deepfake technology following several widely publicised incidents involving manipulated videos of public figures and celebrities. These incidents demonstrated the ease with which synthetic media can be produced and shared across digital platforms, raising concerns about the misuse of AI for harassment, misinformation, and identity manipulation. The widespread availability of generative AI

applications has further increased the risk that similar techniques may be used against ordinary citizens, particularly women, who often possess fewer resources to seek legal remedies or restore their online reputation.

The legal implications of AI-generated deepfakes extend beyond individual privacy violations. They challenge fundamental rights relating to dignity, equality, freedom of expression, and informational privacy while simultaneously testing the effectiveness of existing cyber laws. The anonymous and transnational nature of online platforms further complicates investigation and prosecution, as perpetrators frequently exploit encrypted communication channels, fake digital identities, and foreign-based hosting services. Consequently, traditional approaches to cybercrime enforcement may prove inadequate when addressing AI-enabled offences.

Addressing cyber violence against women therefore requires a multidisciplinary response involving law, technology, public policy, and digital ethics. Effective regulation should combine clear statutory provisions, rapid content-removal mechanisms, improved digital forensic capabilities, platform accountability, victim-support services, and public awareness initiatives. Such an approach would strengthen the protection of women's digital rights while encouraging the responsible development and deployment of Artificial Intelligence technologies.

6. Indian Legal Framework Governing AI-Generated Deepfakes

The increasing misuse of AI-generated deepfakes has exposed significant challenges for India's legal system. Although India does not currently have a dedicated statute regulating deepfake technology, several existing laws provide partial remedies against offences involving synthetic media. These legal provisions collectively address issues relating to cybercrime, privacy, obscenity, identity theft, cheating, defamation, and data protection. However, the rapid evolution of generative Artificial Intelligence has highlighted the limitations of laws that were enacted before the emergence of sophisticated AI technologies.

6.1 Information Technology Act, 2000

The Information Technology Act, 2000 remains the principal legislation governing cyber offences in India. Although the Act does not specifically mention deepfakes or generative Artificial Intelligence, several of its provisions may be invoked against individuals responsible for creating or distributing manipulated digital content.

Section 66C criminalises identity theft involving the fraudulent use of another person's electronic signature, password, or unique identification information. In cases where deepfake creators unlawfully impersonate victims using their digital identity, this provision may become applicable. Similarly, Section 66D penalises cheating by personation using computer resources and communication devices, making it relevant where deepfakes are used to deceive individuals for financial gain or other unlawful purposes (Information Technology Act, 2000).

Section 67 prohibits the publication or transmission of obscene electronic material, while Sections 67A and 67B prescribe enhanced penalties for sexually explicit content and child sexual abuse material circulated through electronic means. These provisions may be invoked where AI-generated intimate deepfakes are created or disseminated without the victim's consent. However, because the legislation predates generative AI, it does not expressly recognise synthetic media or establish liability for AI developers and service providers.

6.2 Bharatiya Nyaya Sanhita, 2023

The Bharatiya Nyaya Sanhita (BNS), 2023, which replaced significant portions of the Indian Penal Code, introduces modernised criminal law

provisions that may indirectly address offences involving AI-generated deepfakes. Acts involving identity impersonation, forgery, criminal intimidation, defamation, sexual harassment, voyeurism, extortion, and publication of false information may be prosecuted under the relevant provisions of the BNS depending upon the factual circumstances of each case.

Where deepfake technology is used to create sexually explicit material targeting women, offences relating to outraging modesty, stalking, criminal intimidation, and sexual harassment may also become applicable. Nevertheless, the BNS does not specifically define or regulate synthetic media generated through Artificial Intelligence, leaving significant interpretative challenges for investigators and courts.

6.3 Digital Personal Data Protection Act, 2023

The enactment of the Digital Personal Data Protection Act, 2023 (DPDP Act) represents an important step towards strengthening individual control over personal data in India. Since the creation of deepfakes frequently involves the unauthorised collection and processing of facial images, voice samples, and biometric information, the Act assumes increasing significance in protecting informational privacy.

The DPDP Act establishes obligations relating to lawful processing of personal data, consent, purpose limitation, and accountability of data fiduciaries. Although the legislation is not specifically designed to regulate AI-generated deepfakes, its emphasis on lawful data processing may provide an important legal basis for addressing unauthorised use of personal information in developing synthetic media (Digital Personal Data Protection Act, 2023).

6.4 Intermediary Liability and Platform Responsibility

Digital platforms play a central role in the creation, dissemination, and amplification of deepfake content. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 impose due diligence obligations upon intermediaries, requiring them to remove unlawful content upon receiving valid legal orders or user complaints. Social media platforms are expected to implement grievance redressal mechanisms, cooperate with law enforcement agencies, and exercise reasonable diligence to prevent the misuse of their services.

Despite these obligations, practical challenges remain. The rapid spread of manipulated content, encrypted messaging applications, anonymous accounts, and cross-border hosting frequently hinder timely detection and removal of harmful material. Consequently, questions regarding the extent of intermediary liability and proactive content moderation continue to generate legal debate.

6.5 Constitutional Protection

The misuse of AI-generated deepfakes directly affects several constitutional rights guaranteed under the Constitution of India. The right to privacy recognised by the Supreme Court in Justice K.S. Puttaswamy (Retd.) v. Union of India forms an essential constitutional safeguard against unauthorised collection, manipulation, and dissemination of personal information (Justice K.S. Puttaswamy (Retd.) v. Union of India, 2017). Similarly, Articles 14 and 21 protect equality before law, dignity, and personal liberty, all of which are undermined when individuals become victims of AI-enabled cyber abuse.

At the same time, regulatory responses to deepfakes must remain consistent with the constitutional guarantee of freedom of speech and expression under Article 19(1)(a). Any restrictions on AI-generated content should therefore satisfy the constitutional standard of reasonableness under Article 19(2), ensuring an appropriate balance between technological innovation and the protection of individual rights.

6.6 Critical Evaluation

Although India's legal framework provides multiple statutory remedies for offences associated with AI-generated deepfakes, these provisions operate in a fragmented manner. Existing laws primarily address the consequences of harmful conduct rather than the unique technological characteristics of generative AI. There is no comprehensive legislation defining deepfakes, allocating liability among AI developers and digital platforms, prescribing standards for watermarking or content authentication, or establishing specialised investigative mechanisms. Consequently, the existing legal framework requires substantial reform to effectively address the emerging challenges posed by synthetic media and AI-enabled cyber violence against women.

7. Comparative Legal Perspectives on Regulating AI-Generated Deepfakes

The rapid proliferation of AI-generated deepfakes has prompted several countries and international organisations to develop legal and regulatory frameworks aimed at mitigating the risks associated with synthetic media. Although approaches differ across jurisdictions, there is growing consensus that existing cybercrime and privacy laws alone are insufficient to address the challenges posed by generative Artificial Intelligence. A comparative examination of international legal responses provides valuable insights for strengthening India's regulatory framework.

7.1 European Union

The European Union has adopted one of the most comprehensive regulatory approaches towards Artificial Intelligence through the Artificial Intelligence Act (AI Act). The Act establishes a risk-based regulatory framework under which AI systems are classified according to the level of risk they pose to individuals and society. While the AI Act does not prohibit deepfake technology, it imposes transparency obligations requiring AI-generated or manipulated content to be clearly disclosed, except in limited situations involving law enforcement, public security, or artistic expression (European Parliament & Council of the European Union, 2024).

In addition to the AI Act, the Digital Services Act strengthens the accountability of online platforms by requiring them to address illegal content, improve transparency, and implement mechanisms for reporting harmful material. Together, these legislative measures seek to balance technological innovation with the protection of fundamental rights, consumer safety, and democratic integrity.

7.2 United States

The United States has not enacted a comprehensive federal law specifically regulating deepfakes. Instead, several states have introduced legislation targeting specific forms of AI-generated synthetic media. For example, states such as California, Texas, and Virginia have enacted laws addressing election-related deepfakes, non-consensual sexually explicit deepfakes, and online image-based abuse.

The regulatory approach in the United States generally emphasises civil liability, criminal sanctions for specific offences, and platform moderation policies. While these state-level initiatives provide protection in particular contexts, the absence of a uniform federal framework has resulted in variations in legal standards and enforcement mechanisms across jurisdictions (Citron & Chesney, 2019).

7.3 United Kingdom

The United Kingdom has adopted a broader online safety strategy to address emerging digital harms. The Online Safety Act, 2023 imposes statutory duties on online platforms to reduce harmful content and strengthen user protection. The legislation requires service providers to implement effective risk management systems,

remove illegal material, and cooperate with regulatory authorities.

Recent legal reforms have also strengthened protections against image-based sexual abuse by criminalising the sharing and creation of intimate images without consent. Although these provisions were not originally designed for AI-generated content, they provide an important legal basis for addressing non-consensual intimate deepfakes targeting women.

7.4 Lessons for India

The comparative analysis demonstrates that several jurisdictions are moving beyond traditional cybercrime legislation by introducing AI-specific regulatory measures, transparency requirements, and enhanced platform accountability. Unlike the European Union, India currently lacks a dedicated legal framework governing the creation, dissemination, and detection of AI-generated deepfakes.

India can draw several lessons from international practices. First, statutory recognition of deepfakes as a distinct category of synthetic media would improve legal certainty. Second, mandatory labelling or watermarking of AI-generated content could reduce misinformation and facilitate digital verification. Third, stronger obligations should be imposed on social media platforms and AI service providers to detect, report, and promptly remove malicious deepfake content. Finally, specialised digital forensic infrastructure and international cooperation should be strengthened to improve the investigation and prosecution of cross-border cyber offences involving generative AI.

While legal responses must remain flexible enough to encourage technological innovation, they should also ensure adequate protection of privacy, dignity, equality, and freedom from online abuse. A balanced regulatory framework incorporating transparency, accountability, and effective enforcement mechanisms would significantly strengthen India's capacity to combat AI-enabled cyber violence against women.

8. Challenges in Investigation and Enforcement of AI-Generated Deepfake Offences

Despite the availability of legal provisions to address cybercrime, the investigation and prosecution of offences involving AI-generated deepfakes remain particularly challenging. The rapid advancement of generative Artificial Intelligence has enabled the creation of highly realistic synthetic media that is often difficult to distinguish from authentic digital content. Consequently, law enforcement agencies, forensic experts, policymakers, and judicial institutions face significant technical, legal, and procedural obstacles in responding effectively to deepfake-related crimes.

8.1 Detection and Authentication of Deepfakes

One of the primary challenges lies in identifying whether digital content has been manipulated using Artificial Intelligence. Modern generative AI models produce highly convincing images, videos, and audio recordings with minimal visual inconsistencies, making manual detection increasingly unreliable. Although AI-based detection tools have been developed, they often struggle to keep pace with the rapid evolution of deepfake generation techniques, creating an ongoing technological competition between content creators and detection systems (Chesney & Citron, 2019).

8.2 Attribution of Criminal Liability

Establishing the identity of individuals responsible for creating or disseminating malicious deepfakes presents another significant challenge. Perpetrators frequently employ virtual private networks (VPNs), encrypted communication platforms, anonymous accounts, and foreign-based servers to conceal their identities. The transnational nature of cyberspace often complicates investigations, particularly where evidence is stored outside India's

jurisdiction or where offenders operate from multiple countries. Such circumstances require effective international cooperation and mutual legal assistance mechanisms.

8.3 Challenges Relating to Digital Evidence

The admissibility and authenticity of digital evidence remain crucial concerns in deepfake investigations. Investigating agencies must establish that electronic records have not been altered and satisfy the procedural requirements governing electronic evidence. Since deepfake technology itself manipulates digital content, distinguishing genuine evidence from fabricated material becomes increasingly difficult. Maintaining the integrity of electronic evidence, preserving metadata, and ensuring a secure chain of custody are therefore essential for successful prosecution (Goodfellow et al., 2014).

8.4 Platform Accountability and Content Moderation

Social media platforms and digital intermediaries play a central role in the dissemination of AI-generated content. Although many platforms have introduced policies for reporting manipulated media, harmful deepfakes often spread rapidly before they can be detected or removed. Delays in content moderation may significantly increase the harm suffered by victims. Furthermore, the absence of uniform standards regarding AI-generated content creates inconsistencies in platform responses and enforcement practices.

8.5 Impact on Victims

Victims of AI-generated deepfakes frequently experience severe emotional, psychological, and social consequences. The circulation of fabricated intimate images or videos may lead to anxiety, depression, reputational damage, workplace discrimination, educational disruption, and social isolation. Women are particularly vulnerable because deepfake technology is often used to create sexually explicit content without consent. Fear of public stigma, victim-blaming, and lengthy legal proceedings may discourage victims from reporting offences, resulting in substantial underreporting of AI-enabled cyber abuse (Citron, 2014).

8.6 Capacity Constraints within Law Enforcement

The investigation of AI-generated cybercrime requires specialised knowledge in digital forensics, cybersecurity, and Artificial Intelligence. However, many investigating agencies continue to face shortages of trained personnel, advanced forensic laboratories, and technological infrastructure. Continuous capacity-building programmes, specialised cybercrime units, and interdisciplinary collaboration between legal experts and technology professionals are essential to improve the effectiveness of investigations involving generative AI.

8.7 Need for International Cooperation

Deepfake-related offences frequently involve multiple jurisdictions because digital content can be created in one country, hosted in another, and accessed globally within seconds. Effective regulation therefore depends upon international cooperation in evidence sharing, extradition, cyber intelligence, and harmonisation of legal standards. Strengthening collaboration among governments, international organisations, technology companies, and law enforcement agencies will be essential for combating cross-border AI-enabled cybercrime.

The challenges discussed above demonstrate that legal reform alone cannot effectively address AI-generated deepfake offences. A comprehensive response requires technological innovation, institutional capacity building, international cooperation, responsible platform governance, and increased public awareness. Without coordinated action, the misuse of generative Artificial Intelligence is likely to continue posing

serious risks to women's digital rights, cybersecurity, and public trust in digital communication.

9. Policy Recommendations

The growing misuse of AI-generated deepfakes against women calls for a comprehensive legal and institutional response that balances technological innovation with the protection of fundamental rights. Since existing legal provisions were not specifically designed to regulate generative Artificial Intelligence, India must adopt a proactive and multidisciplinary strategy to address emerging cyber threats.

9.1 Enact a Dedicated Deepfake Regulation

India should introduce a dedicated legal framework specifically addressing AI-generated synthetic media. Such legislation should define deepfakes, distinguish lawful and unlawful uses, prescribe civil and criminal liability, and establish penalties for the malicious creation or dissemination of deepfake content. A clear statutory definition would reduce ambiguity and facilitate more effective investigation and prosecution.

9.2 Strengthen Platform Accountability

Social media companies, search engines, and AI service providers should be subject to enhanced due diligence obligations. Platforms should implement AI-assisted detection systems, establish expedited grievance redressal mechanisms, preserve digital evidence upon receiving complaints, and remove unlawful deepfake content within a reasonable time. Transparency reports detailing the number of deepfake complaints received and resolved should also be published periodically to promote accountability.

9.3 Develop Advanced Digital Forensic Infrastructure

The Government of India should invest in specialised digital forensic laboratories equipped with AI-based detection tools capable of identifying manipulated audio, video, and image content. Law enforcement agencies should receive continuous technical training to keep pace with evolving generative AI technologies. Collaboration between forensic scientists, cybersecurity experts, and legal professionals should be strengthened to improve investigative efficiency.

9.4 Improve Victim Protection Mechanisms

Victims of AI-generated cyber abuse often face psychological trauma, social stigma, and financial loss. Therefore, legal reforms should incorporate victim-centric remedies, including expedited takedown procedures, confidential reporting mechanisms, legal aid, counselling services, and compensation where appropriate. Dedicated cyber support centres should be established to provide immediate assistance to victims of image-based abuse and identity manipulation.

9.5 Promote Ethical AI Governance

AI developers should adopt ethical design principles that minimise opportunities for misuse. Technical safeguards such as digital watermarking, content authentication technologies, provenance tracking, and disclosure labels for AI-generated content can improve transparency and reduce the spread of manipulated media. Industry self-regulation, combined with statutory oversight, would encourage responsible innovation while maintaining public confidence.

9.6 Enhance Public Awareness and Digital Literacy

Public education remains one of the most effective tools for preventing cybercrime. Educational institutions, government agencies, and civil society organisations should conduct awareness campaigns explaining the risks

associated with deepfakes, methods for verifying digital content, available legal remedies, and safe online practices. Special emphasis should be placed on empowering women and young internet users to recognise and report AI-enabled cyber abuse.

9.7 Strengthen International Cooperation

Given the cross-border nature of cybercrime, India should actively participate in international initiatives relating to Artificial Intelligence governance, cybercrime investigation, and digital evidence sharing. Greater cooperation with foreign governments, international organisations, and technology companies would improve information exchange, facilitate extradition where appropriate, and strengthen the investigation of offences involving transnational digital platforms.

9.8 Encourage Interdisciplinary Research

Universities, research institutions, and policymakers should collaborate to study the long-term legal, ethical, technological, and societal implications of generative Artificial Intelligence. Evidence-based policymaking supported by interdisciplinary research will enable the legal system to respond more effectively to future technological developments.

10. Conclusion

Artificial Intelligence has transformed digital innovation and offers significant opportunities for economic development, scientific advancement, and improved public services. However, the emergence of AI-generated deepfakes has simultaneously created new challenges for cybersecurity, privacy, and the protection of human rights. The increasing use of synthetic media to facilitate cyber violence against women demonstrates how rapidly evolving technologies can be exploited for unlawful purposes when appropriate legal safeguards are absent.

This study examined the technological foundations of deepfakes, their growing misuse against women, and the adequacy of the existing Indian legal framework in addressing such offences. The analysis revealed that although the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Digital Personal Data Protection Act, 2023, and related regulatory measures provide remedies for certain cyber offences, they do not comprehensively regulate AI-generated synthetic media. The absence of explicit statutory provisions, evolving technological capabilities, jurisdictional complexities, and practical difficulties in digital investigation continue to limit effective enforcement.

The comparative analysis further demonstrated that several jurisdictions have begun adopting AI-specific regulatory measures, transparency obligations, and enhanced platform accountability. These developments indicate a global recognition that conventional cyber laws require adaptation to address the unique risks posed by generative Artificial Intelligence.

Addressing AI-enabled cyber violence against women requires more than legislative reform alone. Effective governance must integrate legal regulation, technological innovation, institutional capacity building, ethical AI development, platform responsibility, digital literacy, and international cooperation. Such a multidisciplinary approach is essential to safeguard privacy, dignity, equality, and access to justice while fostering responsible innovation.

As AI technology advances further, the law needs to be dynamic and proactive. It is possible for India to formulate a proper set of regulations that safeguard basic human rights without hindering the advancement of technologies. With proper legal reforms and proper institution building, it can ensure that AI technology is used as a medium of social empowerment and not as a

means of cyber victimization. In essence, protection against AI-based cyber abuse of women is both a technological problem and a constitutional one.

REFERENCES

1. Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review, 107(6), 1753–1820.
2. Citron, D. K. (2014). Hate Crimes in Cyberspace. Harvard University Press.
3. Europol. (2024). Facing Reality? Law Enforcement and the Challenge of Deepfakes. https://www.europol.europa.eu
4. McGlynn, C., Rackley, E., & Houghton, R. (2017). Beyond "Revenge Porn": The Continuum of Image-Based Sexual Abuse. Feminist Legal Studies, 25(1), 25–46.
5. Ministry of Electronics and Information Technology. (2023). Digital Personal Data Protection Act, 2023. Government of India.
6. Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
7. Jain, S. N. (2016). Legal Research and Methodology (8th ed.). LexisNexis.
8. Europol. (2024). Facing Reality? Law Enforcement and the Challenge of Deepfakes.
9. UNESCO. (2023). Guidance for the Governance of Digital Platforms.
10. World Economic Forum. (2024). Global Risks Report 2024.
11. Jain, S. N. (2016). *Legal Research and Methodology* (8th ed.). LexisNexis.
12. Digital Personal Data Protection Act, 2023 (India).
13. Information Technology Act, 2000 (India).
14. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.
15. Justice K. S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1.
16. Citron, D. K., & Chesney, R. (2019). Deep Fakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics. Foreign Affairs, 98(1), 147–155.
17. European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).
18. Online Safety Act 2023 (UK).
19. Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review, 107(6), 1753–1820.
20. Citron, D. K. (2014). Hate Crimes in Cyberspace. Harvard University Press.
21. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems, 27, 2672–2680.
22. Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review, 107(6), 1753–1820.
23. Citron, D. K. (2014). Hate Crimes in Cyberspace. Harvard University Press.
24. Digital Personal Data Protection Act, 2023 (India).
25. European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).
26. Europol. (2024). Facing Reality? Law Enforcement and the Challenge of Deepfakes.
27. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial

Nets. Advances in Neural Information Processing Systems, 27, 2672–2680.
28. Information Technology Act, 2000 (India).
29. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (India).
30. Jain, S. N. (2016). Legal Research and Methodology (8th ed.). LexisNexis.
31. Justice K. S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1.
32. McGlynn, C., Rackley, E., & Houghton, R. (2017). Beyond "Revenge Porn": The Continuum of Image-Based Sexual Abuse. Feminist Legal Studies, 25(1), 25–46.
33. Ministry of Electronics and Information Technology. (2023). Digital Personal Data Protection Act, 2023. Government of India.
34. Online Safety Act 2023 (UK).
35. Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
36. UNESCO. (2023). Guidance for the Governance of Digital Platforms.
37. World Economic Forum. (2024). Global Risks Report 2024.