Page: 1146-1177 *URL: www.lijdlr.com*

LAWFOYER INTERNATIONAL JOURNAL OF DOCTRINAL LEGAL RESEARCH

ISSN: 2583-7753

LawFoyer International Journal of Doctrinal Legal Research

[ISSN: 2583-7753]

Volume 4 | Issue 1

2026

DOI: https://doi.org/10.70183/lijdlr.2026.v04.50

PLATFORM LIABILITY AND DEEPFAKE PORNOGRAPHY: ARE INDIA'S INTERMEDIARY RULES FIT FOR THE AI AGE?

Saloni Shashank Patil[1] & Satya Prakash Mishra[2]

I. ABSTRACT

The advent of deepfake pornography significantly changes the way digital sexual abuse can be carried out. Deepfake pornography is a method of identity distortion and a violation of the individuals' dignity, privacy, and sexual autonomy rights, which are protected by Article 21 of the Indian Constitution. It poses as a serious challenge to the constitutional adequacy of India's intermediary liability framework. "The present regulatory framework under Section 79 of the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 rests on assumptions of intermediary neutrality and notice-based takedown, reflecting a pre-AI model of online harm. This article maintains that the existing framework, which is primarily a reactive takedown one, is constitutionally inadequate in a context where algorithmic systems are used not only to fabricate but also rapidly amplify synthetic sexual content. Based on constitutional jurisprudence recognising dignity and privacy as integral components of the right to life and personal liberty under Article 21 of the Constitution, the paper argues that the State has a corresponding duty to recalibrate platform governance beyond mere passive safe-harbour protection.[3] Thus, the article advocates that instead of reacting to a notice-and-takedown system, a dignity-centric framework should be implemented, which, among other things, recognises the structural responsibility of platforms for AI-enabled harm. The article employs doctrinal analysis of Indian constitutional jurisprudence and a comparative examination of emerging regulatory approaches particularly the European Union's Digital Services Act and the United Kingdom's Online Safety Act to demonstrate how algorithmic amplification undermines the legal fiction of platform neutrality. The article contributes to debates on digital

[1] Chhatrapati Shivaji Maharaj University, Panvel, Navi Mumbai, Maharashtra, India, Email ID: salonispatil123@gmail.com
[2] Chhatrapati Shivaji Maharaj University, Panvel, Navi Mumbai, Maharashtra, India, Email ID: drspm11@gmail.com
[3] *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 S.C.C. 1 (India).

governance and platform accountability by reinterpreting intermediary liability through a constitutional dignity framework in the age of artificial intelligence.

II. KEYWORDS

Deepfake Pornography, Intermediary Liability, Platform Accountability, Algorithmic Amplification, Constitutional Dignity.

III. INTRODUCTION

The increasing reliance on AI-powered content creation has greatly altered the landscape of online harmful content. Among the most damaging effects has been the emergence of deepfake porn. The introduction of deepfake porn is a groundbreaking change in the sphere of digital sexual abuse. Deepfake porn is a synthetic sexual content where a person's face is digitally superimposed in explicit content without their permission.[4] It separates the idea of consent from creation, thus enabling sexual exploitation even if no prior intimate material is present. Deepfake pornography differs from other forms of internet abuse as it not only distributes illegal content, but it actually creates sexual violations through automated methods, making identity, body, and sexuality vulnerable to algorithmic manipulation.

The harm inflicted is often immediate and difficult to reverse, particularly because synthetic sexual content can be replicated and redistributed across multiple platforms within minutes, leaving victims with limited practical avenues for containment.[5] Traditionally, the victims only become aware of the incident when their images or videos have already been spread extensively and further amplified by the algorithm of platforms that are designed to maximize engagement and virality.[6] The damage is much more than just a bruised reputation; it entails deep psychological wounds, alienation from society, and irrevocable loss of one's sexual autonomy.[7] After the leak,

[4] Human Rights Council, Special Rapporteur on Violence Against Women, Its Causes and Consequences, *Report on Online Violence Against Women and Girls from a Human Rights Perspective*, ¶¶ 34–36, U.N. Doc. A/HRC/38/47 (June 18, 2018).
[5] Robert Chesney & Danielle K. Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 Calif. L. Rev. 1753 (2019).
[6] Centre for Countering Digital Hate, *Artificial Amplification: How Social Media Algorithms Promote Sexual Abuse Content* 7–12 (2021).
[7] CLARE MCGLYNN ET AL., SHATTERING LIVES AND MYTHS: A REPORT ON IMAGE-BASED SEXUAL ABUSE 4–6 (2019).

it is almost impossible to get the content removed since it will continue to show up on various platforms, search engines, and mirror sites even after numerous takedown requests.

This technological development exposes the limitations of existing legal assumptions governing online platforms. While certain provisions of the Information Technology Act, 2000 such as Section 66C (identity theft), Section 66D (cheating by personation using computer resources), Section 66E (violation of privacy), and Section 67A (publication or transmission of sexually explicit material) address aspects of digital impersonation, privacy violations, and sexually explicit content, these provisions were not designed to address the algorithmic creation and rapid dissemination of synthetic sexual media.

Consequently, Indian intermediary regulation continues to operate on frameworks that presume passive facilitation and reactive harm. It believes that identifiable human actors generate unlawful content, and it can be effectively controlled through post-publication removal.[8] Deepfake pornography totally shatters this presumption. It is created in large quantities, spread instantly, and kept going by the algorithmic recommendation systems that operate without human intent. In this scenario, regarding platforms as neutral channels is neither an accurate portrayal nor is it normatively defensible.[9]

Accordingly, a critical question arises as to whether the existing legal regime in India adequately protects dignity, privacy, and sexual autonomy as guaranteed under Article 21 of the Constitution. The problem is that India's intermediary liability regime is based on a notice-and-takedown system. The paper contends that India's intermediary framework for liability is, by its nature, incapable of dealing with deepfake pornography unless it undergoes a dignity-focused adjustment. Deepfake porn, or the use of a person's likeness in a doctored sexual video, involves algorithmic

[8] Information Technology Act, 2000, No. 21, Acts of Parliament, 2000, § 79 (India); Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1, ¶¶ 117–18 (India); Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3 (India).
[9] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), arts. 26–27, 2022 O.J. (L 277) 1.

 (ISSN: 2583-7753)

creation and the quick spreading of synthetic sexual content. The intermediary framework is founded upon Section 79 of the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, which are still heavily reliant on conventional norms and understandings, such as the presumption of human wrongdoing and platform neutrality.

A. Research Objectives

1. To assess whether India's intermediary liability laws are sufficient in dealing with the damages caused by deepfake porn.
2. To analyse the constitutional implications of platform regulation from the perspective of the individual's dignity, privacy and sexual autonomy as guaranteed under Article 21 of the Indian Constitution.
3. To assess the boundaries of intermediaries' neutrality and develop a dignity-based model of platform responsibility.

B. Research Questions

1. Is intermediary liability framework of India robust enough to counter the deleterious effects of deepfake pornography?
2. Can we really say that intermediaries that help share information online are neutral (principle of intermediary neutrality) when it comes to content that is made by artificial intelligence and spread by algorithms?
3. To what extent does the intermediary regulation help the State do its job of upholding the constitutional duties as it must do under Article 21 to protect the dignity, privacy and sexual autonomy of individuals?
4. What changes are needed to be made to the law to shape a dignity-centric platform liability framework in India?

C. Research Hypotheses

1. The existing intermediary framework in India, which is based on notice-and-takedown, is structurally insufficient to effectively deal with AI-generated sexual abuse.

2. Algorithmic amplification disrupts the conventional intermediary neutrality doctrine that is at the core of the safe harbour regime.
3. The interpretation of Article 21 from the perspective of human dignity would require digital platforms to have additional proactive obligations.

D. Research Methodology

This work employs qualitative doctrinal research methods primarily based on constitutional and statutory analysis. The study examines judicial decisions on intermediary liability and fundamental rights jurisprudence to evaluate the adequacy of the existing legal framework. In addition to the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, the analysis also acknowledges the relevance of the Digital Personal Data Protection Act, 2023, particularly in relation to questions of consent, personal data processing, and identity-based harms that arise in the context of deepfake pornography.

The methodological approach is anchored in a human rights framework that emphasises dignity, privacy, and sexual autonomy as normative foundations for digital governance. While doctrinal analysis enables a structured examination of constitutional principles and statutory frameworks, it also presents certain limitations. In particular, the harms associated with deepfake pornography involve empirical, technological, and socio-technical dimensions such as algorithmic amplification and platform design which cannot be fully captured through purely doctrinal analysis. The study also adopts a limited comparative approach by examining regulatory developments in the European Union and the United Kingdom in order to identify emerging standards of platform accountability and risk-based regulation.

E. Literature Review

Existing studies concerning digital sexual abuse and platform accountability indicate that there is a gradual transformation of the regulatory focus from content to a more comprehensive analysis of the structural harms ingrained in platform architectures. Initial legal scholarship on the distribution of non-consensual intimate images

concentrated mainly on breaches of privacy and the responsibilities of individuals, thus advocating for the enhancement of criminalization and victim-centric remedies within the existing legal frameworks.[10] Later works extended this conversation by identifying sexual privacy as an essential part of one's autonomy and dignity and highlighting the insufficiency of traditional obscenity-based measures in dealing with technologically-facilitated harms.[11] Although this body of work establishes a normative framework for understanding digital sexual abuse, it generally does not account for synthetic media technologies and therefore does not fully address the algorithmic dimensions of deepfake pornography.

Scholars such as Tarleton Gillespie and Julie E. Cohen have challenged the assumption that intermediaries operate merely as passive conduits, emphasizing that algorithmic curation, moderation practices[12], and data-driven governance significantly shape the visibility and risk environment of online content.[13] Thus, recent literature has focused on the research of platform roles in the development of digital speech environments. Also, these works argue against the law fiction of the intermediary neutrality by saying that the platform delves into the design decisions to the extent that they become the cause of the proliferation of the most harmful content.

However, much of this scholarship has developed within Western regulatory debates and does not directly engage with the constitutional framework of India or the statutory intermediary regime under the Information Technology Act, 2000.[14] While Indian legal scholars have examined intermediary liability and digital governance such as Gautam Bhatia's work on constitutional free speech and digital regulation and Apar Gupta's scholarship on intermediary liability and online speech there remains

[10] Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345, 350–55 (2014).
[11] Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1876–80 (2019).
[12] Julie E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford Univ. Press 2019).
[13] Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (Yale Univ. Press 2018).
[14] Apar Gupta, *Intermediary Liability and Free Speech in India*, various writings and policy commentary on digital rights and platform regulation.

limited academic engagement with deepfake pornography specifically from the perspective of constitutional dignity and platform accountability.[15]

Robert Chesney and Danielle Keats Citron identify deepfakes as a significant threat to privacy, democratic discourse, and personal dignity, arguing that such technologies require regulatory responses that move beyond reactive takedown models toward more comprehensive legal safeguards.[16] Hence, the extant literature on artificial intelligence (AI) and synthetic media technologies have woven a tapestry of complexities for the old rules that were made before the internet. If we look at how platforms are governed, especially in the European Union and the United Kingdom we can see that closer attention is being paid to the systemic risks and algorithmic accountability, making sure that the computer programs used are fair and responsible which is a big part of regulating the internet.[17]

However, the issue of deepfake pornography has so far not been thoroughly explored by Indian legal scholars from the perspective of constitutional dignity and the positive obligation of the State under Article 21. This article, therefore, intends to fill the gap by combining the doctrine of intermediary liability with a constitutionally based framework focusing on human dignity. The article thus adds to the discussions about platform accountability and digital governance in the AI era.

IV. DEEPFAKE PORNOGRAPHY AND THE POLITICAL ECONOMY OF PLATFORMS

Deepfake porn represents a significant evolution in the area of digital sexual abuse. Deepfake porn is the algorithmic generation of sexual content, and hence, it is completely different from the traditional non-consensual intimate imagery, which basically refers to the sharing and exploitation of existing photos or videos without the victim's permission.[18] Machine learning methods are applied to superimpose a

[15] Gautam Bhatia, *Offend, Shock, or Disturb: Free Speech under the Indian Constitution* (Oxford Univ. Press 2016).
[16] Robert Chesney & Danielle K. Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 Calif. L. Rev. 1753 (2019).
[17] Digital Services Act, 2022 O.J. (L 277) at 1; Online Safety Act 2023, c. 50 (U.K.).
[18] Chesney & Citron, *supra* note 9, at 1760–62.

person's face, voice, or body onto pornographic content to create a fake, but highly believable, impression of a sexual act. The violation of the victim's rights occurs not only at the stage of distribution but also at the moment of creation, since the individual's likeness is digitally manipulated without consent. For this reason, deepfake pornography constitutes a form of synthetic sexual violation.

Deepfake content, once made, can be copied, changed with very little 'token' effort, and shared on different platforms within minutes, endlessly.[19] The pornographic deepfake legal issues are not only focused on its being a lie but also on its vastly scalable and permanent nature. Deepfakes, in contrast to other defamatory or obscene materials, exploit the visual credibility of audio-visual media, creating what scholars have described as a form of 'misleading authenticity' that can intensify reputational and dignitary harm.[20] The victim is left with no other choice but to be in a defensive mode-he is explaining, denying, and trying to justify at the same time, the digital imprints are spreading uncontrollably. Therefore, the damage is not an isolated event but a deep one, as it is embedded in the very nature of the technology of communication through the platform.

This kind of structural damage can't be imagined without exploring the role of online platforms. Social media companies, content, hosting sites, and also messaging apps are not simply passive channels through which information is transmitted in a neutral way.[21] They are complex socio-technical systems that have been designed to curate, rank, and amplify content in the pursuit of engagement-driven business strategies. Thus, algorithmic recommendation systems amplify attention-driven material that causes controversy and stirs up emotional traits that deepfake porn usually has.[22] Consequently, the spread of such material is not accidental but rather the direct result of platform design decisions that algorithmically privilege virality over truthfulness and speed over thoroughness.

---

[19] Report on Online Violence Against Women and Girls, *supra* note 1, ¶¶ 34–38.
[20] Robert Chesney & Danielle K. Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 Calif. L. Rev. 1753 (2019).
[21] Gillespie, *supra* note 8, at 21–25.
[22] SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER 82–85 (2019).

Despite this fact, Indian intermediary law still operates on a narrow understanding of platform neutrality concept. Section 79 safe harbour provisions of the Information Technology Act, 2000 depict a straightforward linear model of harm, where the sequence is unlawful content being uploaded, reported, and removed, thereby the liability being triggered only if the actual knowledge is established.[23] Focusing almost exclusively on content removal, the law completely ignores and even rejects the impact of algorithmic distribution. Therefore, this law imposes a disproportionate burden on the victims and, at the same time, provides a shield for the platforms behind which they can remain unaccountable for the systemic harm.

This reactive model is far from sufficient for the time when sexual abuse material is AI-generated, since the harm arises basically from the algorithmic amplification, before any notice can be issued or processed.[24] Hence, deepfake pornography is an example of how a misalignment of the use of the concepts in biometric AI and a foundational point of the Indian intermediary liability framework still exists at a deep level. Deepfake pornography brings out the fact that platforms not merely act as intermediaries between speakers and audiences but have turned into active spaces where harm is created, amplified, and even normalized.[25]

The debate is not centred on the question of whether platforms are the actual producers of deepfake pornographic content, but rather if their technological frameworks, along with regulatory protections, are facilitators of its massive spreading. Therefore, any major legal strategy for deepfake pornography enforcement can no longer be confined to the debate of publisher versus intermediary. It must also come to terms with the platform governance issues. In the subsequent parts of the paper, it will be argued that India's current intermediary liability framework is inadequate in this aspect.

Through sticking to a reactive and content-specific regulatory model, it overlooks the systemic risks of algorithmic amplification and thus, effectively renounces the

[23] The Information Technology Act, 2000, § 79.
[24] *Shreya Singhal*, (2015) 5 S.C.C. 1, ¶¶ 117–19.
[25] Cohen, *supra* note 8, at 147–50.

responsibility of protecting victims of digital sexual abuse in the age of artificial intelligence, which is here to stay.

V. LIABILITY OF INTERMEDIARIES UNDER INDIAN LAW: A STATUTORY AND JUDICIAL ANALYSIS

The main way to control what intermediaries do in India is through Section 79 of the Information Technology Act from 2000. This act works with the Information Technology rules from 2021 which are also known as the Intermediary Guidelines and Digital Media Ethics Code.[26] When these enactments are construed collectively, they create a safe harbour structure that allows intermediaries to be insulated from the liability for third-party content upon fulfilment of the stipulated due diligence conditions.[27] In simple terms it means that intermediaries will not be held responsible for things that other people post if they follow the steps and meet the necessary requirements to be careful. While this regime was intended to maintain equilibrium between the freedom of expression, innovation, and accountability, its basic premises render it incapable of dealing with deepfake pornography issues as well as other forms of AI-generated sexual exploitation.

A. Section 79 of the Information Technology Act, 2000

Section 79 of the Information Technology Act, 2000 establishes the statutory safe harbour regime for intermediaries. However, the present formulation of this provision is the result of a significant legislative revision introduced through the Information Technology (Amendment) Act, 2008. This amendment was enacted in the aftermath of the litigation in *Avnish Bajaj v. State (NCT of Delhi)* and sought to clarify the conditions under which intermediaries could claim immunity from liability for third-party content. Under the amended framework, an intermediary is not held liable for information, data, or communication links hosted by it, provided that it acts merely as a facilitator and does not initiate, select, or modify the transmission of such content, while also observing the prescribed due diligence obligations.

[26] The Information Technology Act, 2000, § 79; IT Rules, 2021, r. 3.
[27] The Information Technology Act, 2000, § 79(1)–(2).

The intermediary is seen as a neutral channel, distinct from the content it hosts, and the liability arises only when such neutrality is compromised through active participation or deliberate ignorance after receiving a notice.[28] The provision makes the granting of immunity subject to the intermediary complying with the standard of care and immediately removing the unlawful content once it has acquired actual knowledge of the fact that the content is illegal. This model is based on the premise that intermediaries do not have much control over the distribution of content before notification and that the damage can be remedied effectively through intervention after publication.

Dissemination of synthetic sexual content is frequently driven not by single uploads but by algorithmic systems that both curate and amplify content based on engagement metrics. Therefore, deepfake pornography is fundamentally against these theories. The impact of recommendation engines, trending algorithms, and automated content ranking systems is that they play a significant role in determining the visibility of the content; however, Section 79 does not specify if these functions constitute the 'selection' or 'modification' of content.[29] By failing to take into account algorithmic curation, the law continues to make a formal distinction between facilitation and participation, which is no longer in line with the present technological realities.

B. Interpreting Intermediary Immunity: Judicial Perspectives

In *Shreya Singhal v. Union of India*, the Supreme Court clarified that intermediaries are required to act only upon receiving 'actual knowledge' of unlawful content through a court order or notification from an appropriate government authority. In the same judgment, the Court also struck down Section 66A of the Information Technology Act, 2000 as unconstitutional for violating the freedom of speech guaranteed under Article 19(1)(a) of the Constitution. The Court's reasoning in *Shreya Singhal* therefore served a dual function: it removed an overly broad speech restriction while simultaneously clarifying the contours of intermediary liability under Section 79 by limiting the circumstances in which platforms could be compelled to remove online content.

[28] *Shreya Singhal*, (2015) 5 S.C.C. 1, ¶¶ 112–17.
[29] The Information Technology Act, 2000, § 79 (containing no reference to algorithmic ranking, recommendation systems, or automated amplification).

Hence, the Supreme Court's reading of Section 79 has cemented the concept of intermediaries as mere passives. These interpretations were made to limit intermediary liability cases and to avoid excessive censorship and a chilling effect on free expression. While Shreya Singhal is an important judgment that upholds digital free speech, the reasoning is aligned with a pre-AI era view of online harms. The Court's stipulation for a formal notice as a condition for the removal of content assumes that illegal content can be identified, assessed, and removed before it is able to cause any irreversible harm.[30]

In the domain of deepfake pornography, this assumption is incorrect. Formal notice and removal actions may be too late after the content has already been copied to different platforms, saved on third, party sites, and indexed by search engines, thus removal will only be a symbolic action. Judicial interpretation of intermediary liability has also evolved through subsequent litigation connected to the *Avnish Bajaj* case. In *Sharat Babu Digumarti v. Government (NCT of Delhi)*, the Supreme Court clarified the relationship between Section 79 of the Information Technology Act and offences under the Indian Penal Code, particularly Section 292 dealing with obscene material.

The Court held that once the conduct in question falls within the specialised framework of the IT Act, prosecution under the IPC for the same conduct would generally be inappropriate. This decision reinforced the IT Act's status as a specialised statutory regime governing online intermediary liability. However, subsequent legislative changes as well as court clarifications have lessened this recognition; thus, the liability model now principally focuses on the observance of the procedures rather than on the avoidance of substantive harm. Overall, the entire legal framework results in a situation where intermediaries are very largely shielded from liabilities even in cases where their technological systems are indeed aiding in the spreading of harmful content.

[30] Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 423–25 (2017).

C. The Intermediary Guidelines and Digital Media Ethics Code Rules, 2021

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 were initially introduced to operationalise the due diligence obligations under Section 79 and to enhance procedural accountability for online intermediaries. However, the regulatory framework has since undergone significant evolution through subsequent amendments. Most notably, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2025, and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 introduced a number of provisions directly addressing synthetic and AI-generated content.

The 2026 Amendment Rules introduce a statutory definition of "Synthetically Generated Information (SGI)" under Rule 2(1) (wa), which expressly encompasses deepfake content. They further require intermediaries enabling the creation or dissemination of such material to implement proactive technical safeguards, including labelling mechanisms for AI-generated content and accelerated response systems for harmful synthetic media. Significantly, the amendments introduce a three-hour takedown requirement for certain categories of SGI-related harmful content, including non-consensual intimate imagery, and prohibit the generation or dissemination of synthetically generated sexual material without consent.

These developments indicate a gradual shift in Indian intermediary regulation from a purely reactive notice-and-takedown model toward a more proactive risk-based framework addressing AI-enabled harms.

Such requirements include appointing grievance officers, issuing user policies and removing illegal content without delay once notified. Although the recent amendments indicate a movement toward proactive regulation of synthetic media, significant questions remain regarding the adequacy, implementation, and constitutional compatibility of these measures. The continuing challenge lies in determining whether these regulatory developments sufficiently address the structural harms associated with algorithmic amplification and AI-generated sexual abuse.

For victims of deepfake pornography, the Rules hardly offer any real support. The requirements of removing pieces of content are triggered only once a complaint has been made, hence the victims are entirely burdened with the detection as well as reporting. The Rules do not require intermediaries to take the initiative in using technology, based safeguards, e.g., deepfake detection tools, although the availability of such technologies. Even though the Rules set out the removal of certain types of content, e.g., sexually explicit material through impersonation within specified timeframes, the implementation of these prerequisites is non, uniform and depends on the discretion of the platforms.

Intermediaries are under no obligation to assess or counteract the overall effect of algorithmic amplification, and neither is there any recognition of platform liability for repeated or viral distribution. Significantly, the Rules still define harm in terms of individual pieces of content rather than considering systemic risks. The emphasis of the regulation is still on the compliance process rather than on delivering positive results for the victims.

D. The Entrenchment of Platform Neutrality Narratives

Even though there has been a rise in public recognition of algorithmic governance, the law still treats intermediaries as neutral technological entities, whose only reaction should be to report illegality once notified. So, a common thread seen repeatedly in India's legislation on intermediary liability is the continuation of platform neutrality as a legal concept.[31] Such a position basically provides an impunity blanket to the platforms for the foreseeable negative effects of their design decisions.

When the law treats intermediary liability in black-and-white terms, i.e., either as a publisher or a passive conduit, it shuts out more nuanced responsibility models that would consider different degrees of control and influence. In that respect, deepfake porn is a strong example of how platforms can, at the same time, refuse to be considered the author, but, by their actions, have a large impact on the distribution. This absence of a legal framework dealing with the platforms' mixed role results in

[31] Gillespie, *supra* note 8, 21–24.

the victims being significantly short of the remedy, and at the same time, the continuation of systemic harms.

E. Rethinking the Regulation of Digital Sexual Abuse

The legal and judicial connotations of intermediary liability in India at first glance reveal a staggering regulatory mismatch. The current system, which prioritises safe harbour protections at the expense of proactive accountability, in effect, sacrifices the dignity of victims to platform immunity.[32] In fact, laws created for an earlier digital era are being stretched to cover AI-related harm that was never imagined.

The above section revealed that India's intermediary liability framework is largely based on obsolete understandings of neutrality, notice, and harm. The next section will show how these doctrinal limitations translate into real institutional failures when deepfake pornography is considered a type of digital sexual abuse. Therefore, the problem is not just about the law being enforced; it calls for a fundamental rethinking of the legal framework that governs platform accountability in the age of artificial intelligence.

VI. REGULATORY AND INSTITUTIONAL GAPS IN TACKLING DEEPFAKE PORNOGRAPHY

India's intermediary liability laws have historically struggled to address the challenges posed by deepfake pornography and, in certain respects, have allowed such harms to persist within existing regulatory structures. The growing institutional concern surrounding deepfake harms is reflected in developments before the Delhi High Court. In November 2024, while hearing a public interest litigation concerning deepfake regulation, the Court directed the Central Government to constitute a committee to examine the regulatory challenges posed by deepfakes, review comparative international frameworks, and undertake consultations with relevant stakeholders.

This judicial intervention highlights the increasing recognition of deepfake-related harms as a pressing governance issue requiring coordinated regulatory responses. The

[32] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1, ¶¶ 297–99 (India).

focus on procedural compliance, the facade of platform neutrality, combined with the overburdening of victims under the current framework basically, results in a mechanism that is legally correct but unresponsive to AI-based sexual abuse at the core.[33] So it's more than just enforcement gaps or technical limitations that have resulted in the failure; the failure is actually a feature of the very setting of the rules.

A. Victim-Centred Enforcement and Regulatory Displacement

The framework of intermediaries' notice-and-takedown system operates on the assumption that victims are best positioned to identify, report, and seek removal of unlawful content. While the Information Technology Act remains the primary statutory framework governing intermediary liability, the broader criminal law architecture has also undergone significant transformation with the enactment of the Bharatiya Nyaya Sanhita, 2023, which replaced the Indian Penal Code. Certain provisions of the BNS may become relevant in addressing coordinated deepfake pornography campaigns.

For instance, Section 353, which penalises the dissemination of false or misleading information, and Section 111, which addresses organised criminal activity including cyber-enabled offences, may potentially be invoked in cases involving systematic production or distribution of synthetic sexual content. But in many deepfake pornographic cases, victims do not discover the synthetic sexual content until it has become highly popular, being shared across numerous platforms, private groups, and even archival websites.[34] So, when a notice is finally sent, the harm has usually been done to a great extent. So, the above assumption does not work in the case of deepfake pornography. Even when victims come forward, the implementation process remains fragmented and ambiguous.[35]

Each platform has different reporting mechanisms; there is no regular enforcement of timelines, and the final decisions on removals still depend on the discretion of the platforms. The law does not rank sexual deepfakes as a priority, even though they

[33] Citron, *supra* note 7, at 1902–05.
[34] Clare McGlynn et al., *supra* note 3, at 81–83.
[35] Report on Online Violence Against Women and Girls, *supra* note 1, ¶¶ 44–47.

inflict harm that is very severe and long-lasting in a unique way.[36] Therefore, this victim-centred enforcement model functions more as a house of cards than a source of power; it makes individuals, who are in a highly vulnerable situation, carry the load of responsibility that should be shared by platforms and regulators.

B. Algorithmic Amplification and Regulatory Inaction

Deepfake pornography overwhelms the internet not as a result of random occurrence, but because it is deliberately circulated via recommendation systems targeting user engagement.[37]

Therefore, the primary law instrument to oversee the regulation of the issue is found in the way it deals with algorithmic amplification. The popularity, spread, and copying of such content are mainly due to intentional platform design decisions; they are not the result of a neutral process.

Nevertheless, on the matter of the legal implications of algorithmic promotion, Indian intermediary law still remains significantly silent. Section 79, as well as the 2021 Rules, do not require the evaluation, modification, or accountability of recommendation systems in any circumstance.[38] Amplification is considered as a matter of technical necessity rather than a legally significant conduct which leads substantially to harm. This ambiguity enables the platforms to shift the blame for the consequences they proverbially summon to nature and thereby hides systemic risks under claims of technological neutrality.[39]

Merging hosting and amplification into one category of facilitation, the law blurs distinctions that are essential to understanding platform power. In the age of AI, this mixing up is impossible to maintain. If platforms develop a system that forecasts, highlights, and commercializes the most harmful content, it is practically impossible for them to continue being legally recognized as mere passive intermediaries.

---

[36] IT Rules, 2021, pt. II sec. 3(i), r. 3(2).
[37] Gillespie, *supra* note 8, at 62–67.
[38] The Information Technology Act, 2000, § 79.
[39] FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 141–44 (2015).

C. Procedural Compliance at the Expense of Substantive Protection

India's intermediary regulations lay more emphasis on procedural compliance rather than on the accomplishment of substantive results. As accountability indicators, the number of grievance officers, the publishing of user policies, and adherence to takedown schedules are taken into account.[40] However, such formal conditions are hardly effective in stopping the abuse or assisting the victims in a significant way after the abuse has taken place. Removal of content duties can only happen once a complaint is filed, thus ignoring the sudden harm caused by deepfake pornography.[41] Grievance procedures operate in a non-transparent manner; also, there are no binding requirements to judge their effectiveness. Compliance is gauged more on procedures than on the results, thus allowing platforms to satisfy the regulatory requirements while the essential aspects of harm remain the same.

This procedural emphasis also breaks down systemic abuse into a series of individual acts. For instance, deepfake porn may be used as one of the ways in a continuous harassment or extortion campaign; however, the regulatory framework treats each posting as a separate violation. The law, therefore, by not dealing with the harm caused by the perpetrator and the repetitions, enables the platforms to deal with the superficial issues while the root causes remain unrecognized and unrectified.

D. Regulatory Omission of Preventive and Positive Obligations

Platforms are not obliged to prevent the risks that may be caused by their technologies, even if they have modes of detection and protective measures at their disposal.[42] Hence, a major and serious flaw of India's intermediary framework is the absence of enforceable preventive duties. This issue is particularly worrying as deepfake pornography is expected to increase in digital spaces. The hesitance in the formulation of positive obligations reveals a regulatory attitude that is more inclined to innovation than to safety. By the law's figurative acceptance of harm as a tolerable externality, it

[40] *Shreya Singhal*, (2015) 5 S.C.C. 1, ¶¶ 116–17.
[41] Chesney & Citron, *supra* note 9, in 1763–66.
[42] Eur. Parl. Directorate-Gen. for Internal Policies, The Impact of Algorithms for Online Content Filtering or Moderation 18–20 (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/657101/IPOL_STU(2020)657101_EN.pdf, [visited on 26-02-2026].

is basically normalizing the occurrence of digital sexual abuse instead of dealing with it.[43] However, a prioritization of this kind raises constitutional issues when technological systems are used to subject individuals to serious violations of dignity and sexual autonomy.

E. Institutional Failure and the Normalisation of Harm

Taken together, these regulatory limitations reveal a broader pattern of institutional fragmentation. While courts, legislatures, and regulatory authorities have begun acknowledging the governance challenges posed by synthetic media, the legal framework remains distributed across intermediary regulation, criminal law, and emerging policy initiatives. The absence of clear coordination between these institutional responses contributes to uncertainty in enforcement and leaves victims navigating a fragmented regulatory landscape. Victims are urged to depend on legal and regulatory systems that assure protection, but oftentimes result in delays, deflections, and procedural fatigue. Platforms invoke safe harbour provisions, regulators stress formal compliance, and the State withdraws behind private governance frameworks.

When deepfake pornography is regarded as an unfortunate yet unavoidable aspect of technological advancement, the law conveys that dignity is negotiable in online environments. The long-term effect is the "normalisation of digital sexual harm". This degradation of normative commitment erodes public trust, as well as diminishes the credibility of constitutional safeguards in digital realms. This section has demonstrated that India's intermediary liability framework not only fails to preclude deepfake pornography; it also perpetuates conditions that allow such harm to thrive. Addressing this issue necessitates more than enhanced enforcement it requires a reconfiguration of platform accountability that acknowledges algorithmic influence, accentuates victim dignity, and enforces substantial preventive responsibilities.

[43] *Puttaswamy*, (2017) 10 S.C.C. 1, ¶¶ 119–121.

VII. RETHINKING PLATFORM RESPONSIBILITY THROUGH COMPARATIVE REGULATION

So far, there is no legal system that has tried to solve the problems of deepfake pornography in India. Thus, India's framework on intermediary liability appears to be less and less adequate as one compares it to the changing regulatory policies of other regions. However, the latest developments in the *European Union* and the *United Kingdom* show a marked change from the traditional, notice, based regulation to the models that recognize the platform authority and systemic risk.[44] These strategies highlight the conceptual limits of India's continuing reliance on intermediary neutrality.

A. The United Kingdom: Preventive Obligations and Platform Accountability

The *Online Safety Act* of the United Kingdom has brought about a huge change by setting up legal duties for online platforms.[45] Contrast the intermediary framework in India, which sees harm as incidental and reactive, the UK approach emphasizes prevention and risk mitigation. The platforms have to make reasonable and proportionate efforts to prevent the spread of illegal content. Also, they have to be more responsible for content that is related to sexual exploitation and abuse.[46] Platforms are expected to put in place mechanisms and procedures, including the use of automated tools, to spot and eliminate risks even before the harm occurs.[47] I It is noteworthy that the Act recognises that certain categories of harm require rapid and proactive intervention.

At the same time, the Online Safety Act 2023 has been the subject of significant debate regarding its potential chilling effect on freedom of expression. Critics have argued that expansive preventive obligations and the threat of regulatory penalties may incentivise platforms to engage in overly cautious moderation practices that could restrict lawful speech. Nevertheless, the legislation represents a clear normative shift:

[44] Digital Services Act, 2022 O.J. (L 277) at 1; Online Safety Act 2023, c. 50.
[45] Online Safety Act 2023, c. 50, §§ 1–4.
[46] Id. §§ 9–12; U.K. Dep't for Digital, Culture, Media & Sport, *Online Safety Act: Explanatory Notes* (2023).
[47] Online Safety Act 2023, c. 50, §§ 14–18.

platform immunity is no longer treated as automatic but is increasingly conditioned upon demonstrable efforts by intermediaries to prevent foreseeable harm."

B. The European Union: Prioritizing Systemic Risk over Isolated Content

The Digital Services Act (DSA) of the European Union introduces differentiated responsibilities for platforms based on their size, reach, and capacity to influence the dissemination of information, shifting the regulatory focus from isolated unlawful content toward systemic risk management. Complementing this framework, the European Union has also adopted the Artificial Intelligence Act (Regulation (EU) 2024/1689), which introduces transparency obligations for AI-generated content and imposes specific regulatory requirements on certain categories of high-risk AI systems. Among other provisions, the AI Act requires clear disclosure when content has been synthetically generated or manipulated, including deepfakes, thereby addressing the informational asymmetries that allow synthetic media to mislead audiences. Thus, by shifting the regulatory emphasis towards systemic risk, the DSA marks a major departure from traditional intermediary liability laws.[48] It has put in place obligations for *Very Large Online Platforms* (VLOPs) to identify and mitigate systemic risks that may result from the dissemination of illegal and harmful content, such as those risks created by algorithmic recommendation systems.[49]

Deepfake pornography is one type of content that this approach would be highly effective at solving. To some extent, the DSA recognizes that the main source of harm is not always the quantity of content but the factors that make it more visible. Regulatory instruments such as the Digital Services Act and the EU AI Act collectively demonstrate an emerging European approach that links platform governance with algorithmic transparency, risk assessment, and disclosure obligations for AI-generated content. These measures reflect a recognition that harms associated with synthetic media arise not only from individual pieces of content but also from the technological infrastructures that enable their creation and amplification. The DSA does not mention deepfake pornography specifically as a problem. Still, its emphasis

[48] Digital Services Act arts. 26–28, 2022 O.J. (L 277) at 1.
[49] Digital Services Act art. 34, 2022 O.J. (L 277) at 1; European Commission, *The Digital Services Act Package* (2022).

on systemic risk reduction offers a regulatory model that is missing in Indian law, which is still very much focused on post-hoc content removal only.

C. The Limits of Regulatory Transplantation: The Indian Experience

Before committing to any one model, differences in constitutional settings, regulatory powers, and market structures must be thoroughly weighed.[50] So, such comparative methods are undoubtedly subject to debate and criticism, and they cannot be applied to India straightforwardly. The regulatory frameworks of both the EU and the UK have identified that platform governance should deal with aspects such as algorithmic design, amplification, and systemic risk components that are largely absent from the intermediary liability framework in India.[51] Hence, the thrust of the comparison lies not in the duplication of their measures but in the differentiation of one's measures.

By making amplification legally insignificant and prevention merely discretional, Indian laws fall short of the capability of adjusting to the global standards that are gradually recognizing platforms as being significantly responsible for the digital harm ecosystems. Hence, it is clear that India has been mostly taking the straightforward ways of doing things for a very long time. India has, for a long time, been dependent on a minimalist complaint-based approach only. India has been sticking to this method for so long that it has become redundant. It clearly shows that the people-in-charge are still thinking like they did when the internet first started. Thus, it leads to real-life situations where victims of deepfake pornography are unable to get the protection they deserve, as neither the people nor the legal regime is changing with the times. This happens simply because the law is more inclined to accept the authority of the platform rather than challenge it. India is required to change its present legal framework completely if it is to address the challenges of the AI era. Also, India needs these changes to help people who are affected by deepfake-enabled abuse.

[50] Anupam Chander & Uyen P. Lê, *Data Nationalism*, 64 EMORY L.J. 677, 721–24 (2015).
[51] *Shreya Singhal*, (2015) 5 S.C.C. 1; The Information Technology Act, 2000, § 79.

D. Normative Perspectives on Indian Law

Currently, platforms hold a *hybrid role*, exerting considerable influence over visibility and reach while denying authorship.[52] So, the comparative experience provides a distinct normative perspective: "the efficacious regulation of deepfake pornography necessitates moving beyond the dichotomy of publisher versus passive intermediary."[53] Regulatory frameworks that do not recognize this hybridity are fundamentally unable to tackle AI-enabled sexual abuse.

Consequently, India's intermediary liability framework does not indicate a lack of alternatives, but rather a hesitance to reconceptualize responsibility.[54] As will be discussed in the following section, selectively drawing from comparative insights, especially the acknowledgement of systemic risk and preventive duty, presents a pathway toward a regulatory framework that harmonizes platform regulation with constitutional dignity and the protection of victims.

VIII. RETHINKING PLATFORM LIABILITY IN THE AGE OF ARTIFICIAL INTELLIGENCE

Thus, the previous analysis clearly shows that the pattern of India's intermediary liability laws is largely out of touch with the nature of AI-based digital sexual abuse. Addressing deepfake porn should not mean a total discard of safe harbour protections. It should be more of an adjustment that considers platform power, algorithmic governance, and constitutional duties. As a result, a reimagined platform liability framework should not only be about reactive compliance but also move to a regulatory approach that is human dignity-centred and harm-based.

A. Beyond Reactive Takedowns: Embedding Proactive Risk Responsibility

Deepfake pornography is something that we should have seen coming. It is not an unusual exception anymore. This is a normal result of people using artificial intelligence in a legally impermissible manner on the internet and other digital systems. So, we shouldn't be surprised anymore. Deepfake pornography is one of the

[52] Pasquale, *supra* note 35, at 59–63.
[53] Gillespie, *supra* note 8, at 45–48.
[54] DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 167–70 (Harvard University Press 2014).

things that happens when people misuse artificial intelligence within modern digital ecosystems. Consequently, the foundation of reform should involve a transition from notice-centric regulation to risk-based governance. Platforms utilizing generative and recommendation technologies are in a unique position to foretell and alleviate such risks. Consequently, intermediary liability ought to be contingent not only on prompt responses to complaints but also on clear efforts to avert the spread of such synthetic sexual content.

This transition requires the establishment of proportionate preventive responsibilities tailored to the capabilities and reach of each platform. It does not suggest imposing strict liability on platforms. Instead, it mandates major platforms to implement appropriate safeguards, such as content detection tools and watermarking systems, along with friction-based measures for categories of high-risk content. It also urges such major platforms equipped with advanced algorithmic systems to perform regular risk assessments that evaluate the potential for deepfake misuse. The fact that Indian law does not have rules for this now means we have a big problem that needs to be fixed immediately. It signifies a regulatory void that needs urgent attention.

B. Algorithmic Amplification and Its Legal Implications

When platforms actively promote content through algorithms optimized for engagement, they wield a type of curatorial power that necessitates appropriate legal accountability. Therefore, a major flaw in India's intermediary framework is clearly evident: It simply cannot recognize algorithmic amplification as a legally significant type of conduct. It fails to understand that when algorithms make something more visible, that is a kind of action that should be considered by law. Intermediaries can say they are not responsible for things that nobody sees or that are not amplified. They may justifiably assert immunity. But this does not make sense when platforms use algorithms to make harmful content more visible. Also, recommendation systems are not impartial technical processes; rather, they represent design choices that significantly affect visibility, reach, and potential harm. These choices can really affect how many people see something and how much harm it can cause.

India's intermediary framework needs to be reformed so that it can deal with amplification in a better way. Introducing higher due diligence requirements in cases of amplified spread would be a way of bringing legal accountability in line with technological realities. The reform must espouse a nuanced approach to liability that differentiates between mere hosting and active amplification. This way, we can make sure that intermediaries are not treated like publishers.

C. Fast-Track and Victim-Oriented Takedown Procedures

It is quintessential for regulatory reform to focus on speed and accessibility for victims. It is especially needed in cases involving sexual content that features impersonation or synthetic alterations. This is due to the fact that deepfake pornography inflicts such extensive harm at a rate that traditional takedown procedures are simply not able to cope with. The law needs to accept that victims shouldn't have to prove that the video is fake or that the person who made it did so on purpose. It would put additional burden on those who are impacted if they are tasked with the burden of proving falsity or intent during the takedown process. So, it's also very important to make it easier for victims to report and alleviating procedural burden. We need an easy process to remove deepfake pornography. We need an open system, for reporting deepfake pornography that works across all platforms so victims don't have to keep submitting the same information over and over. This process should have rules and deadlines because the harm caused by these videos can't be undone. The system should be easy to use and make it easy to report deepfake pornography.

Thus, reporting mechanisms should be uniform, transparent and very easy to get to. This will really help in cutting down on the need to make multiple submissions across different platforms. In instances where credible claims of synthetic sexual content are presented, platforms should be mandated to respond promptly, with the possibility of subsequent review.

D. Platform Regulation through the Lens of Constitutional Dignity

Deepfake pornography constitutes a serious violation of sexual autonomy and informational privacy, as it involves the unauthorized digital manipulation of an

individual's identity and bodily representation. Such harms implicate the constitutional protections of dignity and privacy recognised by the Supreme Court under Article 21.[55] Thus, any reform in the area of intermediary liability has to be anchored on constitutional principles, primarily the right to freedom from degrading treatment and privacy referred to in Article 21. If laws fail to adequately regulate such a wrong, the State's responsibility to protect fundamental rights from violation by non-state actors is endangered.

Instead of viewing victim harm as a peripheral issue that is less important than innovation and free speech, the law should acknowledge dignity as a quintessential standard. Incorporating constitutional dignity into platform regulation does not insinuate a need for total censorship. It necessitates a change in normative perspective. However, it does require that the protection of bodily and mental integrity is accentuated in instances where technological systems enable sexual exploitation. Hence, intermediary liability is all about making sure these platforms do their part. We should think of intermediary liability not just as a matter of statutory protection but as a way to make sure online platforms follow the constitutional mandates that are meant to protect people's rights in the digital world. In terms of intermediary liability is not just about giving intermediaries a free pass or not; it is about ensuring they help enforce constitutional obligations online.

E. Regulatory Oversight as a Tool for Institutional Accountability

Effective regulation of online platforms in the age of artificial intelligence requires institutional mechanisms capable of ensuring transparency, accountability, and compliance with legal obligations. While existing intermediary regulations impose certain procedural duties, much of the responsibility for interpreting and implementing these obligations continues to rest with the platforms themselves. This arrangement creates regulatory asymmetry, as private entities exercise substantial discretion over compliance while public authorities possess limited insight into the technological systems that shape content visibility and distribution.

[55] *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 S.C.C. 1 (India).

Strengthening institutional accountability therefore necessitates more structured regulatory oversight. One potential approach involves the introduction of independent auditing mechanisms through which large online platforms may be required to periodically assess and disclose the risks associated with their algorithmic systems, including those related to the creation or amplification of harmful synthetic media. Such audits would enable regulators to obtain reliable information about platform governance practices while promoting greater transparency in the design and deployment of automated content systems.

At the same time, regulatory oversight must remain proportionate and consistent with constitutional safeguards for freedom of expression and technological innovation. The objective is not to impose intrusive state control over digital platforms but rather to establish mechanisms that ensure responsible platform governance. Within such a framework, harms arising from deepfake pornography may be recognised as a distinct regulatory concern requiring targeted safeguards, thereby enabling legal institutions to respond more effectively to emerging forms of AI-enabled abuse.

F. Reframing Platform Liability through a Dignity Lens

Deepfake pornography is a problem that we can see coming because of the way technology works and the fact that nobody is doing anything to stop it. A framework becomes constitutionally ineffectual if it accentuates platform immunity to the detriment of dignity risks, even if it adheres to statutory requirements. Deepfake pornography is an issue, and we need to think about how it affects peoples' dignity when we make rules about it. Thus, if we make rules that protect websites and do not care about people's dignity, these rules will not work even if they follow the law.

The problem we are facing is not about whether the digital platforms can sustain, but whether our legal system can handle new emerging kinds of harm. We should be concerned about deepfake pornography. To deal with this problem we need an approach that is proactive, fair and based on the principles of our constitution. This is because deepfake pornography is not simply an unintended consequence of user behaviour. The proposed reforms are neither radical nor unprecedented. Deepfake pornography is an issue that requires a thoughtful and principled solution and these

reforms are a step, in the right direction to address deepfake pornography. They signify a gradual yet quintessential evolution of intermediary liability in light of technological advancements. So, reconceptualising platform liability in the era of AI necessitates the recognition of a fundamental reality: "platforms are no longer *mere passive intermediaries*".

IX. SUGGESTIONS AND RECOMMENDATIONS

In light of the foregoing analysis, a number of policy and regulatory reforms may be considered to address the growing threat posed by deepfake pornography in digital environments. While recent developments in intermediary regulation have begun to recognise synthetic media as a distinct category of online harm, further institutional and legal reforms remain necessary to ensure meaningful protection of dignity, privacy, and sexual autonomy.

1. **Adoption of Risk-Based Platform Governance:** Regulatory frameworks should move beyond purely reactive notice-and-takedown models toward risk-based governance approaches. Intermediaries operating large-scale algorithmic systems should be required to conduct periodic risk assessments evaluating the potential misuse of their platforms for generating or amplifying synthetic sexual content. Such assessments would enable early identification of systemic risks associated with deepfake technologies.
2. **Proactive Technical Safeguards for Synthetic Media:** Platforms facilitating the creation or dissemination of AI-generated media should implement technical safeguards designed to detect and mitigate harmful synthetic content. These may include deepfake detection tools, content authentication systems, watermarking technologies, and friction-based controls that limit the rapid virality of high-risk media content.
3. **Fast-Track Takedown Procedures for Non-Consensual Synthetic Sexual Content:** Given the rapid and irreversible nature of harm caused by deepfake pornography, regulatory frameworks should provide accelerated takedown mechanisms for content involving non-consensual synthetic sexual imagery. Such procedures should include simplified reporting mechanisms, clearly

defined response timelines, and prioritised review processes for victims of digital sexual abuse.

4. **Algorithmic Accountability and Transparency:** The role of algorithmic recommendation systems in amplifying harmful content necessitates greater transparency and accountability. Regulatory frameworks should therefore require large platforms to disclose key information about algorithmic ranking systems and undertake independent audits assessing the impact of such systems on the dissemination of harmful synthetic media.
5. **Strengthening Victim-Centred Remedies:** Legal frameworks addressing deepfake harms should incorporate victim-centred remedies that reduce the procedural burden on affected individuals. This may include the establishment of unified reporting channels, standardised complaint mechanisms across platforms, and legal recognition of synthetic sexual content as a distinct form of digital sexual abuse.
6. **Embedding Constitutional Dignity in Platform Regulation:** Any future reforms to intermediary liability must remain anchored in constitutional principles, particularly the protection of dignity, privacy, and sexual autonomy under Article 21 of the Constitution of India. Platform governance should therefore be evaluated not merely in terms of procedural compliance but also in terms of its ability to prevent serious violations of fundamental rights in digital environments.
7. **Institutional Coordination and Regulatory Oversight:** Addressing the harms associated with deepfake pornography requires coordinated institutional responses involving legislatures, regulatory agencies, and courts. Establishing specialised regulatory oversight mechanisms capable of monitoring emerging AI-enabled harms would strengthen the institutional capacity to respond to evolving technological risks.

X. CONCLUSION

Deepfake pornography illustrates how emerging artificial intelligence technologies can transform digital sexual abuse into a scalable and structurally embedded harm

within platform ecosystems. The analysis in this article demonstrates that India's intermediary liability framework was historically designed around assumptions of passive hosting and reactive notice-based enforcement, assumptions that become increasingly inadequate when harmful content can be algorithmically generated and rapidly amplified across digital networks.

Recent regulatory developments, particularly the amendments to the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules introducing obligations relating to synthetically generated information, represent an important institutional recognition of the governance challenges posed by deepfake technologies. These developments indicate an emerging shift toward proactive regulation, including enhanced due diligence requirements and accelerated response mechanisms for harmful synthetic media. Nevertheless, the effectiveness of these measures will ultimately depend on their implementation, the scope of technological safeguards adopted by platforms, and the extent to which regulatory oversight addresses systemic risks associated with algorithmic amplification.

The central argument of this paper has been that intermediary liability should be reconceptualised through a constitutional lens grounded in dignity, privacy, and sexual autonomy under Article 21 of the Constitution. From this perspective, the regulation of deepfake pornography cannot be treated merely as a problem of unlawful content removal. Rather, it raises broader questions concerning platform governance, institutional accountability, and the positive obligations of the State to protect fundamental rights within digital environments.

As artificial intelligence continues to reshape the architecture of online communication, ensuring meaningful protection against digital sexual abuse will require regulatory approaches that recognise the hybrid role of platforms as both intermediaries and influential curators of online visibility. A dignity-centred framework for platform liability therefore remains essential for aligning intermediary regulation with constitutional principles and for ensuring that fundamental rights retain their normative force within rapidly evolving technological landscapes.

 (ISSN: 2583-7753)

XI. REFERENCES

A. Cases

1. Avnish Bajaj v. State (NCT of Delhi), 150 D.L.T. 769 (Delhi H.C. 2008).
2. Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1 (India).
3. Sharat Babu Digumarti v. Government (NCT of Delhi), (2017) 2 S.C.C. 297 (India).
4. Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1 (India).

B. Indian Statutes & Subordinate Legislation

1. Information Technology Act, 2000, No. 21, Acts of Parliament, 2000 (India).
2. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (India).

C. Foreign & International Legislation

1. Digital Services Act, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services, 2022 O.J. (L 277) 1.Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L 1689) 1.
2. Online Safety Act 2023, c. 50 (U.K.).

D. International & Government Reports

1. Human Rights Council, Special Rapporteur on Violence Against Women, Its Causes and Consequences, *Report on Online Violence Against Women and Girls from a Human Rights Perspective*, U.N. Doc. A/HRC/38/47 (June 18, 2018).
2. Centre for Countering Digital Hate, *Artificial Amplification: How Social Media Algorithms Promote Sexual Abuse Content* (2021).
3. CLARE MCGLYNN ET AL., SHATTERING LIVES AND MYTHS: A REPORT ON IMAGE-BASED SEXUAL ABUSE (2019).

4. European Parliament Directorate-General for Internal Policies, *The Impact of Algorithms for Online Content Filtering or Moderation* (2020).
5. European Commission, *The Digital Services Act Package* (2022).
6. U.K. Dep't for Digital, Culture, Media & Sport, *Online Safety Act 2023: Explanatory Notes* (2023).

E. Books

1. DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE (Harvard Univ. Press 2014).
2. JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM (2019).
3. FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015).
4. TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA (2018).
5. SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (2019).

F. Journal Articles

1. Anupam Chander & Uyen P. Lê, *Data Nationalism*, 64 EMORY L.J. 677 (2015).
2. Robert Chesney & Danielle K. Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753 (2019).
3. Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345 (2014).
4. Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401 (2017).
5. Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870 (2019).
6. Carly Parnitzke Smith & Jennifer J. Freyd, *Institutional Betrayal*, 69 AM. PSYCHOLOGIST 575 (2014).