
INTERNATIONAL JOURNAL OF ADVANCED LEGAL RESEARCH

**BEYOND REMOVAL- RE-ASSESSING INTERMEDIARY LIABILITY
FOR THE PERSISTENT CIRCULATION OF NCII AND DEEPPFAKE
CONTENT AGAINST WOMEN IN INDIA**

- Jaskeerat Khaira¹

ABSTRACT

The increasing use of digital platforms and artificial intelligence has contributed to the rise of non-consensual intimate images (NCII) and deepfake-based abuse, particularly against women. Existing cyber laws and intermediary frameworks primarily focus on the identification, reporting, and removal of harmful content. However, harmful content frequently continues to circulate through private messaging, cloud-sharing services, mirror uploads, and other alternative dissemination channels even after formal takedown. Consequently, existing legal mechanisms appear to concentrate on content removal rather than addressing its continued circulation. This reveals a significant gap in intermediary liability frameworks concerning post-removal recirculation and continuing digital harm. It is comprehensible that mere removal of content and no follow-up causes more harm than good. The paper adopts a doctrinal approach by examining Indian cyber laws, intermediary liability principles, and constitutional protections relating to privacy, dignity, and equality. The paper evaluates whether intermediaries should bear continuing obligations in cases of foreseeable recirculation and proposes reforms aimed at strengthening accountability and protection for victims of gender-based digital violence.

Keywords: Non-Consensual Intimate Images (NCII); AI Deepfakes; Intermediary Liability; Content Recirculation; Cyber Law; Gender-Based Digital Violence; Digital Privacy.

INTRODUCTION

¹ Final-Year B.A LL.B. Student from Punjab College of Law, Chunni Kalan (Affiliated to Punjabi University Patiala)

For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

<https://www.ijalr.in/>

The rapid expansion of digital platforms has significantly increased the circulation of non-consensual intimate images (NCII) and AI-generated intimate deepfakes, creating new challenges for legal regulation. Unlike conventional online content, intimate images can be repeatedly copied, re-uploaded and redistributed across multiple platforms, allowing harm to persist even after the original content has been removed. Existing research further demonstrates that these harms are disproportionately experienced by women due to entrenched gender stereotypes, social stigma and psychological trauma, while advances in artificial intelligence have made the creation and dissemination of intimate deepfakes increasingly accessible.

In India, intermediary liability is primarily governed by Section 79 of the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, which adopt a notice-and-takedown approach by requiring intermediaries to remove unlawful content upon receiving a valid legal notice while preserving safe harbour protection. Although this framework facilitates content removal, it raises important questions regarding its effectiveness in addressing the persistent recirculation of NCII and deepfake content.

Against this backdrop, this paper addresses two interrelated questions: first, whether India's existing intermediary liability framework adequately addresses the continued circulation of NCII and AI-generated deepfake content after formal takedown; and second, whether intermediaries should bear continuing obligations to prevent foreseeable recirculation of such content, particularly where women are disproportionately affected.

The study adopts a **doctrinal and comparative legal research methodology**. It analyses the Information Technology Act, 2000, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 and relevant judicial interpretation, while drawing comparative insights from the European Union's Digital Services Act and contemporary technological measures for repeat-upload detection. On this basis, the paper argues that India's predominantly reactive notice-and-takedown framework should evolve towards a calibrated, victim-centred model that imposes limited continuing obligations on intermediaries to address the foreseeable recirculation of NCII and AI-generated deepfakes

For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

<https://www.ijalr.in/>

without undermining constitutional guarantees of free expression and intermediary safe harbour.

THEORETICAL AND REGULATORY FOUNDATIONS OF INTERMEDIARY LIABILITY

The increasing use of social media platforms has created unprecedented opportunities for communication while simultaneously exposing users to growing forms of online abuse. Intermediary liability refers to the legal responsibility of online platforms for unlawful third-party content hosted or transmitted through their services, subject to the statutory immunity available under law. With the rapid expansion of internet accessibility, the tension between freedom of speech and expression and platform accountability has become increasingly significant. Certain posts are shared in the guise of freedom to express their opinion in a democracy, but it often crosses the line of what is legal and what is not. This friction inherently attracts the provisions of intermediary liability. The current legal ecosystem deals with the reactive takedown model. Herein, the model receives a complaint of material uploaded in the present case, NCII and the removal is done either by the Cyber Cell associated or the social media platforms or both. Yet still, after the removal of content through official channels, the same is shared through replicated media, cloud storage, and direct messaging. It not only presents the question of the stringency of the law but also the lack of post-removal safety given to the victim. There arises a need for a proactive mechanism and duty of care for the victim, whether that be refraining the users from sharing the material regardless of the editing and complete omission of material on the platform. This in itself can be a huge coordinating task, but the flagship system and comparative viewpoint can help the liability reach a point. Against this backdrop, this paper examines whether India's existing intermediary liability framework adequately addresses the continued circulation of NCII and AI-generated deepfake content after formal takedown and whether intermediaries should bear continuing obligations to prevent its foreseeable recirculation.

INDIA'S EXISTING INTERMEDIALITY LIABILITY FRAMEWORK: A REACTIVE MODEL

For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

<https://www.ijalr.in/>

Section 79 of the Information Technology Act, 2000² grants “safe harbour” for intermediaries in exceptional cases. It provides a conditional exception where the intermediary is not the primary originator of the illicit material. However, this immunity disappears as soon as the said social media platform engages in or fails to exercise due diligence in the course of action. Herein, by due diligence in accordance with the rules³ we mean publishing updated rules and regulations for the public. Users are to be informed of applicable restrictions as applicable. The content with no owner value may be removed within 36 hours of the complaint received or the order of the Court. It can be clearly deduced that the existing frameworks focus more on the grievance-based removal rather than concentrating on the re-sharing of the content removed.

In the case of *Shreya Singhal v. Union of India*⁴ it was referred by the Hon’ble Court that “actual knowledge” does not arise merely because a private person sends a notice alleging illegality. An intermediary is not expected to determine whether content is lawful or unlawful, as this requires legal adjudication. “Actual knowledge” exists only after the intermediary receives a court order or a lawful direction from the appropriate Government/its agency. Furthermore, the court asks for an order or government intervention for the removal of such content; the intermediary cannot primarily remove content under their autonomy.

Analysing the above provisions and precedent, it is clearly known that the existing system focuses on reactive takedown, that too at the behest of legislation or judiciary. This approach is particularly problematic in cases involving non-consensual intimate images, where the legal framework prioritises initial removal but provides insufficient mechanisms to prevent subsequent re-uploads, mirror copies, and continued circulation, thereby limiting the effectiveness of victim protection.

THE LIMITS OF REACTIVE TAKEDOWNS IN CASES OF NCII AND AI-GENERATED DEEPFAKES

² Information Technology Act, 2000, § 79 (India)

³ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3 (India).

⁴ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1 : AIR 2015 SC 1523 (India)

As noted in the previous report of the Secretary-General, technology-facilitated violence against women and girls takes many forms, including sexual harassment, stalking, zoombombing, image - based abuse, trolling, doxing, misogynistic or gendered hate speech, misinformation and “astroturfing” (the artificial impression of widespread support for an idea), among others. Some forms of violence against women and girls, such as intimate partner or domestic violence and trafficking, are also facilitated through different digital tools, including mobile phones, Global Positioning Systems and tracking devices. The growth of AI is creating new forms of violence against women and girls and new avenues to entrench the normalisation of harmful social norms and violence against women and girls, which will need to be factored into definitions and measurement tools.⁵ The report of the UN General Assembly clearly mentions that the harm so suffered is gendered violence primarily against women and girls. Furthermore, the harms are also gendered by means of the sexualized and misogynistic form and manner in which they are manifested. Without doubt, societal gender disparities, particularly the persistence of sexual double standards, enable humiliation, stigma and shame to be visited on women, facilitating the production and prevalence of image-based sexual abuse.⁶ In the context of understanding of the image-based crime predominantly known as NCII, the method of reactive takedown can fail the victims of justice. The following issues can be addressed to better understand the gap in the present legal frameworks:-

1. Persistent Recirculation Through Mirror Uploads and Cross-Platform Sharing: First and foremost, even after the formal removal of non-consensual sexual content, there follows a cycle of persistent recirculation post-removal. Such recirculation happens in the disguise of edited or modified posts, commenting on the relevant posts, anonymous links or asking for direct messaging(DM), etc. It might also include the domain of Cross-Platform sharing, indicating that certain removal on one platform might not result in hindering any sort of active resharing on other platforms. This asks for cross-platform coordination and active monitoring of even slight changes made to

⁵ UN General Assembly, Report of the Secretary-General on “Advancement of Women: Intensification of Efforts to Eliminate All Forms of Violence against Women and Girls: Technology-Facilitated Violence against Women and Girls”, 79th Session, Agenda Item 27, para 11 at p. 5 (United Nations).

⁶ Clare McGlynn and Erika Rackley, “Image-Based Sexual Abuse” 37 (2017) *Oxford Journal of Legal Studies* 534, 543 (p. 9).

the material being shared. Moreover, keyword detection can be used as an essential tool to track down certain comments or DMs in the app.

2. Continuing Harm and the Burden Placed on Victims: Secondly, the victims at the forefront face the social stigmatisation of being involved in such content, completely ignoring the sensitive nature of their involvement. The usual course of procedure makes it the victim's responsibility to share the complaint, and that results in recurring complaints as the primary content keeps resurfacing through mirror upload, as stated in the above point. This way of dealing with online sexual abuse poses a ton of emotional and social trauma on the victim, with the whole burden bearing of the aftermath regardless of who the offender is.
3. Inadequacy of Existing Obligations in Addressing AI-Generated Deepfakes: Thirdly, as far as the prevailing provisions of law are concerned, we do not have specialised obligations for AI-generated deepfakes. Despite the active development of the law over the course of years, there is no certain provision for morphed material. Even the largely discussed question revolves around who owns the AI-made material, but we fail to acknowledge the identical nature of the content being uploaded to the real identity of the victim. Regardless of the ownership, the focus shall rely upon the abuse and harm faced. The present system requires a clear, distinctive addition of AI sexual abuse, as it differs from general forms of sexual abuse. Even though it might be tempting to add it to the box of cyber sexual harassment but, the terminology would need a separate rulebook to address the fabricated nature of the explicit material shared.
4. Absence of a Continuing Duty of Care: Lastly, the present structure, as discussed above, focuses more on the pre-removal complaint and completely bypasses the care and assistance required after the content has been removed. It clearly demonstrates the absence of mechanisms to address repeat uploads, generate digital hashes of previously identified unlawful content, or adopt reasonable measures to prevent its foreseeable recirculation. Once the reported content is taken down, the intermediary's responsibility largely comes to an end, even though the same material may continue to circulate across multiple accounts and platforms. Consequently, victims are often compelled to repeatedly report identical content, placing an unnecessary burden upon them while allowing the cycle of online abuse to continue. The absence of a

For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

<https://www.ijalr.in/>

continuing duty of care therefore exposes a significant gap in India's intermediary liability framework, particularly in cases involving NCII and AI-generated intimate deepfakes.

The foregoing discussion demonstrates that the existing notice-and-takedown framework is insufficient to address the evolving nature of NCII and AI-generated intimate deepfakes. While the current legal regime provides mechanisms for the removal of unlawful content, it remains largely reactive and fails to account for the persistent recirculation and continuing harm suffered by victims. This gap is further widened by the absence of specialised obligations to address AI-generated abuse and post-removal responsibilities of intermediaries. These shortcomings highlight the need to reconsider intermediary liability beyond mere content removal and towards a framework capable of addressing systemic and recurring forms of online image-based abuse. The following section examines comparative regulatory approaches to identify reforms that may strengthen the Indian legal framework.

COMPARATIVE REGULATORY DEVELOPMENTS: FROM TAKEDOWN TO SYSTEMIC RESPONSIBILITY

Intermediary liability may broadly be understood through two regulatory approaches: content-based regulation and systemic risk regulation. Under a content-based model, the primary obligation of an intermediary is to remove unlawful content after receiving a valid notice, court order, or government direction. This notice-and-takedown approach focuses on individual pieces of illegal content after they have appeared online. By contrast, systemic risk regulation extends platform responsibility beyond content removal by requiring platforms to identify, assess, and mitigate the structural and technological factors that facilitate the dissemination of harmful content.

The *European Union's Digital Services Act (DSA)*⁷ represents a prominent example of this systemic approach. Rather than limiting intermediary obligations to the removal of unlawful content, the DSA imposes additional responsibilities on Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs). Under Article 34, these platforms are required to identify, analyse, and assess the systemic risks arising from the

⁷ Regulation (EU) 2022/2065 (Digital Services Act), 27 October 2022, OJ L 277/1 (European Union)

design, functioning, and use of their services, including risks associated with algorithmic and recommender systems. These assessments must be carried out annually and before introducing new features that are likely to significantly affect such risks.⁸

Article 34 requires platforms to evaluate a broad range of risks. These include the dissemination of illegal content, adverse effects on fundamental rights such as human dignity, privacy, data protection, freedom of expression, and consumer protection, as well as risks to public security and democratic processes. Importantly, the provision expressly recognises risks relating to gender-based violence, the protection of minors, public health, and serious harm to an individual's physical and mental well-being. This explicit recognition is particularly relevant in cases involving non-consensual intimate images (NCII), where victims often experience long-lasting psychological, emotional, and social harm.

The DSA also recognises that online harms are not created solely by users but may be intensified by the way platforms operate. Consequently, Article 34 requires platforms to assess how recommender systems, content moderation mechanisms, advertising systems, and data-processing practices contribute to systemic risks. It further requires consideration of risks arising from automated manipulation, coordinated misuse of services, and the rapid amplification and widespread dissemination of illegal content. Although Article 34 does not expressly refer to the post-removal resharing of non-consensual intimate images, its systemic risk framework is broad enough to encompass recurring harms arising from repeated uploads, algorithmic amplification, and the continued circulation of illegal content.

This approach differs from India's existing intermediary liability framework under Section 79 of the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. The Indian framework primarily relies on notice-and-takedown obligations, under which intermediary responsibility generally arises only after unlawful content has been identified through a complaint, court order, or government notification. While these provisions facilitate the removal of illegal content, they do not require platforms to undertake mandatory assessments of systemic risks or evaluate whether their algorithms, platform design, or operational practices contribute to the spread or

⁸ Regulation (EU) 2022/2065 of the European Parliament and of the Council (Digital Services Act), 27 October 2022, OJ L 277/1, Art. 34 (European Union).

repeated circulation of such content. The DSA therefore adopts a broader preventive model by requiring large platforms to actively identify and reduce systemic risks before they result in widespread harm. In doing so, it extends intermediary responsibility beyond the removal of individual content to addressing the underlying conditions that enable online harms to persist and recur.

TOWARDS A CONTINUING DUTY OF CARE: REIMAGINING INDIA'S INTERMEDIARY LIABILITY

To move forward in the future, we need a robust framework that promotes a continuing duty of care. The following ways can be considered while progressing in the context:

1. Adoption of Hash-Matching and Repeat Upload Detection Technologies: the technological tool of hash matching can be an exceptionally good tool to prevent modified pictures, and it formulates restrictions even before the upload is completed. PhotoDNA creates a unique digital signature (known as a “hash”) of an image, which is then compared against signatures (hashes) of other photos to find copies of the same image. When matched with a database containing hashes of previously identified illegal images, PhotoDNA is an incredible tool to help detect, disrupt and report the distribution of child exploitation material. PhotoDNA is not facial recognition software and cannot be used to identify a person or object in an image. A PhotoDNA hash is not reversible, and therefore cannot be used to recreate an image.⁹ This hash system works based on the similarity ratio of already uploaded or legally provided material. It can help in pre-posting removal as well. Furthermore, in the case of AI-generated deepfake explicit videos, audio fingerprinting can be used as a mechanism. Audio fingerprinting provides the ability to link short, unlabeled snippets of audio content to corresponding data about that content.¹⁰ This sort of method is

⁹ Microsoft Corporation, “PhotoDNA”, available at: <https://www.microsoft.com/en-us/photodna> (last accessed 6 July 2026).

¹⁰ Shumeet Baluja and Michele Covell, “Content Fingerprinting Using Wavelets”, Google, Inc., Technical Report, available at: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/32685.pdf> (accessed 6 July 2026).

For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

<https://www.ijalr.in/>

used where audio files or videos with added audio are concerned. At last, repeated uploads with identical files can also be detected as a method of deterrence.¹¹

2. Establishing a Victim-Centred Reporting and Redressal Mechanism: The existing reporting mechanism places an unnecessary burden on victims by requiring them to repeatedly report the same content whenever it is re-uploaded or shared through different accounts or platforms. This process delays effective relief and forces victims to relive the trauma associated with the abuse. A victim-centred mechanism should therefore allow a single complaint to initiate the removal of all identified copies of the unlawful content. Intermediaries should also provide a simple and accessible reporting system, ensure prompt communication with complainants, and coordinate with other platforms wherever feasible. Such a mechanism would reduce the burden on victims while ensuring quicker and more effective removal of NCII and AI-generated intimate deepfakes.
3. Introducing Specific Obligations for AI-Generated Intimate Deepfakes: The existing legal framework does not specifically address AI-generated intimate deepfakes, despite their increasing misuse for image-based abuse. Although the 2021 Rules recognise morphed images, they do not prescribe separate obligations for detecting, reporting or removing AI-generated explicit content. Given the unique nature of deepfakes, intermediaries should be required to establish dedicated reporting mechanisms and ensure their prompt removal. The law should also provide a clear legal definition of AI-generated intimate deepfakes to avoid uncertainty in enforcement. Introducing specific obligations for such content would strengthen victim protection and ensure that the legal framework keeps pace with emerging technologies.
4. A Continuing Duty of Care for Intermediaries: The present legal framework mainly focuses on removing unlawful content after a complaint is made. However, this is not enough in cases of NCII and AI-generated intimate deepfakes, where the same content can be uploaded again within minutes. Therefore, intermediaries should have a continuing duty of care even after the original content has been removed. They

¹¹ OpenAsset, “How OpenAsset Manages Duplicate Detection”, OpenAsset Help Center, available at: <https://success.openasset.com/en/articles/12674665-how-openasset-manages-duplicate-detection> (accessed 6 July 2026).

should take reasonable steps to prevent the same content from being repeatedly shared, such as using hash-matching technology and detecting repeat uploads. This would reduce the burden on victims and make the existing legal framework more effective. At the same time, these measures should be implemented without affecting the freedom of expression of genuine users.

CONCLUSION

The existing intermediary liability framework under the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, provides an important legal mechanism for the removal of unlawful online content. However, as discussed throughout this paper, it remains largely limited to a reactive notice-and-takedown approach. While such a framework may be effective in removing an individual post, it does not adequately address the repeated circulation of NCII and AI-generated intimate deepfakes through mirror uploads, cross-platform sharing and other forms of re-uploading. As a result, victims continue to face emotional, social and psychological harm even after the original content has been removed.

The comparative approach adopted by the European Union demonstrates that intermediary liability can extend beyond content removal by placing reasonable responsibilities on platforms to identify and reduce recurring online harms. Drawing from these developments, India should strengthen its existing framework by introducing practical measures such as hash-matching technology, victim-centred reporting mechanisms, specific obligations for AI-generated intimate deepfakes and a continuing duty of care to prevent foreseeable re-uploads. These reforms would improve victim protection without disturbing the balance between intermediary immunity and freedom of speech and expression. Ultimately, intermediary liability should not end with the removal of unlawful content but should continue until reasonable steps have been taken to prevent its repeated circulation.

For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

<https://www.ijalr.in/>