PLATFORM ACCOUNTABILITY FOR TECHNOLOGY-FACILITATED GENDER-BASED VIOLENCE IN INDIA: RETHINKING INTERMEDIARY LIABILITY AND CONTENT MODERATION

Lakshmiprasad A.N., LLM, University Law College and Department of Studies in Law, Bengaluru University, Bengaluru, Karnataka

ABSTRACT

Social media and other online platforms have become part of everyday life, but they have also become places where women face harassment, threats, stalking and the sharing of intimate images without consent. This kind of abuse, often called technology-facilitated gender-based violence, does not end when a post is removed. It can push women out of public conversation and affects their privacy, safety and equal participation in society.

This article examines how Indian law deals with the responsibility of platforms for such abuse. It looks at the criminal provisions that apply, the safe harbour for intermediaries under section 79 of the Information Technology Act, 2000, the way the Supreme Court read that provision in Shreya Singhal, and the Intermediary Rules of 2021 with their twenty-four hour takedown rule for intimate imagery. It also considers the Delhi High Court's decision in X v. Union of India and compares the Indian position with the approach of the European Union and the United Kingdom.

The article argues that the present system depends too heavily on victims finding, reporting and following up each piece of abusive content, and that notice and takedown alone is not enough. It suggests that large platforms should be asked to assess and reduce gender-based risks, to make reporting accessible in Indian languages, to deal effectively with copies of intimate imagery and to be transparent about how they respond. It also stresses that these duties must come with safeguards for free speech and privacy.

Keywords: Gender Justice; Technology-Facilitated Violence; Intermediary Liability; Platform Accountability; Content Moderation; Digital Rights.

I. INTRODUCTION

For many women, going online now comes with a kind of background risk. A photograph taken from a social media profile can be edited and posted on an abusive page. A private image shared in a relationship can be uploaded to a website after the relationship ends. A woman who speaks about politics or caste or gender can find her phone number and address circulated, or find that she is receiving threats every time she posts. None of this is new in itself, but the internet makes it faster, cheaper and far harder to undo.

Scholars and international bodies call this technology-facilitated gender-based violence. The label matters because it treats these acts as part of the wider pattern of violence against women and not as an unfortunate side effect of the internet. It also directs attention to the companies that run the platforms where the abuse is posted, shared and sometimes amplified.

India has responded mostly through criminal law and through rules that oblige platforms to remove certain content when they are told about it. These tools are useful, but they work after the harm has already happened. A victim has to find the content, report it, follow it up and often go to the police or a court, while the material may be copied elsewhere in the meantime.

This article asks whether the Indian approach to platform accountability is adequate for gender-based online abuse. It first describes how the abuse works and why it is a digital rights issue as much as a criminal law issue. It then examines the Indian legal framework, with attention to the safe harbour for intermediaries and the Intermediary Rules of 2021. After comparing the position in the European Union and the United Kingdom, it makes some suggestions. The argument is that India needs to move from a system based only on notice and takedown towards one that also asks large platforms to anticipate and reduce gender-based harm, while keeping free speech and privacy safeguards in place.

II. WHAT TECHNOLOGY-FACILITATED GENDER-BASED VIOLENCE LOOKS LIKE

The conduct covered by this term is varied. It includes the sharing of intimate images without consent, which is often called non-consensual intimate imagery, and the creation of morphed or artificially generated sexual images using a woman's face. It includes stalking through messages, tracking apps and repeated contact, and the publication of personal details such as

an address or a workplace in order to invite harassment, a practice known as doxxing. It also includes organised abuse and threats of rape or violence, as well as steady streams of sexist insults aimed at women who are visible in public life.

The harm goes beyond the immediate distress. A woman who has been targeted may delete her accounts, stop speaking in public, leave a profession or become afraid to leave her home. Abuse of this kind pushes women out of public conversation, which means it interferes with their freedom of expression and their equal participation in society. That is why the subject belongs in a discussion of digital rights and not just crime.

International bodies have said as much. The Committee on the Elimination of Discrimination against Women has stated that gender-based violence against women happens in technology-mediated environments as well as in physical spaces, and that states have duties in respect of it.[1] The United Nations Special Rapporteur on violence against women has described online violence as part of a continuum with offline violence and has said that internet intermediaries have a role to play in preventing it.[2]

In India the right to privacy is a fundamental right under Article 21 of the Constitution, as the Supreme Court held in Puttaswamy.[3] A woman's control over images of her body and information about her life falls squarely within that right. Equality under Articles 14 and 15 also matters, because abuse that targets women as a group has the effect of denying them equal access to public life.

III. THE INDIAN LEGAL FRAMEWORK

The first layer is criminal law. The Bharatiya Nyaya Sanhita, 2023, which replaced the Indian Penal Code, deals with several of the relevant acts. It covers sexual harassment, voyeurism, including the capture and sharing of private images, stalking including online stalking, and words, gestures or acts that insult the modesty of a woman.[4]

The second layer is the Information Technology Act, 2000. Section 66E punishes capturing or

[1]Comm. on the Elimination of Discrimination Against Women, General Recommendation No. 35 on Gender-Based Violence Against Women, U.N. Doc. CEDAW/C/GC/35 (July 14, 2017).
[2]Special Rapporteur on Violence Against Women, Report on Online Violence Against Women and Girls from a Human Rights Perspective, U.N. Doc. A/HRC/38/47 (June 18, 2018).
[3]Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1 (India).
[4]Bharatiya Nyaya Sanhita, 2023, No. 45 of 2023, §§ 75, 77, 78, 79 (India).

publishing images of a person's private area without consent, and sections 67 and 67A deal with publishing or transmitting obscene and sexually explicit material in electronic form.[5] These provisions are often used together with the Sanhita in cases of image-based abuse.

The third layer concerns the platforms. Section 79 of the Information Technology Act gives intermediaries a safe harbour from liability for content posted by third parties, provided they meet certain conditions, including observing due diligence.[6] The Intermediary Rules of 2021 set out what due diligence involves. Under Rule 3(2), an intermediary must have a grievance mechanism, and where a complaint is made by an individual or on her behalf about content that exposes her private areas, shows nudity or a sexual act, or impersonates her, including through morphed images, the intermediary must take all reasonable and practicable measures to remove or disable access within twenty-four hours.[7]

The Digital Personal Data Protection Act, 2023 adds a further layer by creating obligations for those who process personal data, and it is relevant where abuse involves the misuse of personal information.[8] Its effect on gendered abuse is still developing.

Taken together, India has more tools than critics sometimes admit. What matters is how they work in practice, and the answer depends heavily on how courts have read the safe harbour.

IV. INTERMEDIARY LIABILITY AND THE LIMITS OF NOTICE AND TAKEDOWN

The leading authority on intermediary liability is the Supreme Court's decision in Shreya Singhal. The Court struck down section 66A of the Information Technology Act and read down section 79(3)(b), holding that an intermediary loses its safe harbour for unlawful content only when it has actual knowledge, which means that it has received a court order or a government notification directing the content to be removed.[9] The purpose was to protect free speech. If platforms had to decide on the legality of every complaint themselves, the court reasoned, they would remove too much to stay safe.

[5]Information Technology Act, 2000, No. 21 of 2000, §§ 66E, 67, 67A (India).
[6]Information Technology Act, *supra* note 5, § 79.
[7]Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, G.S.R. 139(E), r. 3(2) (India).
[8]Digital Personal Data Protection Act, 2023, No. 22 of 2023 (India).
[9]Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1 (India).

That reasoning has real force, but it creates a difficulty for victims of abuse. If a platform is only compelled to act when there is a court order, then a woman whose private images are circulating needs to go to court, and often to do so repeatedly. The Rules of 2021 tried to ease this by creating the twenty-four hour duty for intimate imagery, which does not wait for a court order. In this limited area, then, the law now leans in favour of the victim.

The Delhi High Court's decision in X v. Union of India shows how this works and where it stops. The petitioner's intimate images had been posted on several websites, and although individual links were removed, the images kept being re-uploaded. The court held that intermediaries must take down the offending content itself and not only the specific links given by the victim, and it issued directions on the use of hash-matching technology to detect copies, on sensitising grievance officers and on making the reporting mechanism easy to find.[10] The direction to search engines to identify and de-index duplicate images automatically was challenged in appeal by Google, so its final shape should be checked.[11]

The case shows a basic weakness of notice and takedown. When copies spread across many sites, removing one link at a time resembles a game that the victim cannot win. The problem is not that the law fails to say that takedown is required. It is that the burden of finding and reporting every copy falls on the person who was harmed.

There are genuine counter-concerns. A system that requires platforms to detect and block content automatically may lead to over-removal, may require monitoring of private communication and may make mistakes. These risks affect privacy and speech, including the speech of women themselves, such as survivors who want to talk about their experience or activists who share evidence of abuse. Any reform has to meet this objection and not wave it away.

The Supreme Court has also dealt with the problem in the context of sexual violence content. In the proceedings that began with a letter from the organisation Prajwala, the Court considered how platforms could stop the circulation of videos of sexual violence, and it asked industry and the government to work on workable measures.[12] This shows that the judiciary has already regarded platform responsibility as a legitimate issue and not just a private matter between a

[10]X v. Union of India, 2023 SCC OnLine Del 2361 (India).
[11]Google LLC v. X, LPA 373/2024 (Del. H.C.) (India).
[12]In re: Prajwala Letter, Suo Motu W.P. (Crl.) No. 3 of 2015 (S.C.) (India).

victim and an uploader.

V. CONTENT MODERATION IN PRACTICE

Law on paper only works if platforms apply it well, and here several practical problems arise. The first is access. A victim must be able to find the reporting tool, understand it and use it, often at a time of great distress. If the process is only available in English, or if the forms are long and unclear, many women will not complete them.

The second is language and context. Abuse is frequently written in regional languages, in mixed-language slang or in coded words that automated systems may not recognise. Moderation systems built mainly for English may miss what is plainly abusive to a local reader, or may flag harmless posts. This is an area where platforms with global reach and large resources have the capacity to do better.

The third is design. Platforms do not only host content. Their recommendation systems decide what is shown to whom, and content that provokes strong reactions often travels further. If a platform's design rewards pile-ons and harassing material, then the platform is not a neutral bystander, even if it did not create the content. The law of safe harbour was built around the idea of a passive host, and that picture does not fit modern platforms very well.

The fourth is transparency. At present it is hard for outsiders to know how many complaints about gender-based abuse a platform receives, how quickly it responds, how many are rejected and why. Without that information regulators, courts and researchers cannot judge whether the twenty-four hour rule is being followed in practice.

The fifth is the position of the police and courts. A takedown by a platform does not replace an investigation. Victims often report that officers are untrained in cybercrime or reluctant to take such complaints seriously, and that platforms respond to law enforcement requests slowly. Better cooperation among platforms, police and the courts is therefore as important as the platform's own rules.

VI. LESSONS FROM OTHER JURISDICTIONS

The European Union's Digital Services Act places layered duties on online platforms. All hosting services must have a notice and action mechanism through which people can report

illegal content,[13] and the largest platforms must assess systemic risks arising from their services, including negative effects in relation to gender-based violence, and take reasonable steps to reduce them.[14] This is different from the Indian approach because it asks large platforms to think about risk in advance, instead of reacting only to complaints.

The European Union has also adopted a directive on combating violence against women and domestic violence, which requires member states to criminalise acts such as the non-consensual sharing of intimate or manipulated material, cyber stalking, cyber harassment and cyber incitement to violence or hatred.[15] The United Kingdom's Online Safety Act 2023 likewise places duties on services to deal with illegal content and to assess risks, and it creates offences relating to the sharing of intimate images.[16]

These laws cannot be copied directly. Their obligations are linked to larger regulatory bodies and to conditions in the European legal system, and they have attracted criticism over enforcement and free speech. But they show two ideas that are useful for India. One is that the duty of large platforms need not be limited to responding to a particular complaint. The other is that gender-based violence can be named as a specific category of risk in platform regulation.

VII. CORPORATE RESPONSIBILITY AND HUMAN RIGHTS

Beyond the law of liability there is the idea of corporate responsibility. The United Nations Guiding Principles on Business and Human Rights say that companies have a responsibility to respect human rights, which includes carrying out due diligence to identify, prevent and address the harms their activities may cause.[17] For a social media company this would include harms to women who are abused through its services.

These principles are not binding law, and India has no general statute that makes them so for digital platforms. Still, they provide a useful standard against which platform conduct can be measured, and they are echoed in the risk assessment approach of the European law. In India,

[13]Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act), 2022 O.J. (L 277) 1, art. 16.
[14]Digital Services Act, *supra* note 13, art. 34.
[15]Directive (EU) 2024/1385 of the European Parliament and of the Council of 14 May 2024 on Combating Violence Against Women and Domestic Violence, arts. 5–8.
[16]Online Safety Act 2023, c. 50 (UK).
[17]U.N. Office of the High Commissioner for Human Rights, Guiding Principles on Business and Human Rights, U.N. Doc. HR/PUB/11/04 (2011).

the closest binding obligations are the due diligence rules discussed earlier, which are narrower in scope. A gap therefore remains between the responsibilities that international standards describe and the duties that Indian law enforces.

VIII. SUGGESTIONS FOR INDIA

The discussion above suggests several reforms that could improve platform accountability without giving up the safeguards that Shreya Singhal sought to protect.

First, large platforms should be required to carry out regular assessments of the risks of gender-based abuse on their services and to describe what they are doing about them. This would move part of the burden from the victim to the company, and it can be limited to the biggest platforms so that smaller services are not overloaded.

Second, the process for dealing with intimate imagery should be made more victim-centred. A person should be able to report once and have the platform take reasonable steps to find and remove identical copies, using hash-matching where it is technically suitable, with safeguards on accuracy, on the use of the hashes and on appeal. The platform should also preserve the relevant records so that the police can investigate. The directions given in X v. Union of India are a good starting point.

Third, reporting should be easy to find and available in major Indian languages, and grievance officers should be trained to deal with abuse cases with sensitivity. The time limits in the rules should be backed up by public reporting on how they are met.

Fourth, platforms should publish regular transparency reports that include information on complaints of gender-based abuse, response times and outcomes. This would let regulators and the public check whether the existing rules work.

Fifth, police and judges need training and clear channels for working with platforms. Without it, even the best platform rules will not produce real protection. Digital literacy programmes for women, and for young people in general, can add to this, though they should supplement and not replace the duties of platforms and the state.

Finally, safeguards must be built in. Any obligation to detect content should be narrow, should apply only to clearly defined categories such as non-consensual intimate imagery, should give

users a way to challenge mistakes and should avoid general monitoring of private communications. It would also be better for the main duties to be placed in a statute passed by Parliament rather than left only to rules made under delegated powers, because the rules have already faced challenges and a statute carries more democratic legitimacy.

IX. CONCLUSION

Online abuse of women is not a minor irritation of life on the internet. It affects privacy, safety, speech and equality, and it reduces women's participation in public life. India has a body of criminal and regulatory law that responds to it, and the twenty-four hour takedown rule for intimate imagery and the Delhi High Court's directions show that the law has begun to take platform responsibility seriously.

Even so, the present approach depends too much on victims finding, reporting and following up each piece of abusive content. The safe harbour in section 79 was built for a simpler picture of the internet, and Shreya Singhal rightly protects free speech but leaves a gap where abuse is repeated and rapidly copied.

The way forward is not to remove safe harbour or to require platforms to monitor everything. It is to ask the largest platforms to assess and reduce the risks of gender-based harm, to make reporting simple, to deal with copies of intimate imagery effectively, and to be transparent about how they perform. These measures should come with real protection for free expression and privacy. If that balance is achieved, digital spaces can be made safer for women without closing them to anyone else, and the idea of gender justice can be given practical meaning online.