Cyberviolence Against Women – A Roadmap for Legal Reform

Inputs to the Law Review Consultation Convened by the National Commission for Women

IT for Change

December 2020

Table of Contents

1. Introduction

Sexism, misogyny and gender-based violence on social media is trivialised and dismissed as a normal part of the online experience. Law enforcement agencies and platform intermediaries have failed in evolving effective and timely responses, hampered as they are by pre-digital legal-institutional frameworks whose conceptual frames fall short of grasping the new taxonomies of violence in digital sociality. IT for Change’s survey research with 881 young women college students in Kerala, Karnataka and Tamil Nadu in 2019 demonstrated the ubiquity of sexism, misogyny and gender-based violence in online spaces. Over 3/4th of respondents reported having faced online gender trolling in some form. Bullying based on physical attributes emerges as a common form of violation experienced by young women. 31% of respondents who have faced cyberviolence reported being bullied about their body shape; 30%, their weight; 27%, their looks; and 22%, their skin colour. The repeal of Section 66A of the Information Technology Act (IT Act) leaves women with very little recourse to

pursue cases of gender-trolling. Misogynistic speech or gender-based hate speech is not recognised as aggravated grounds for hate, and the police and lawyers find it difficult within other sections of the Indian Penal Code (IPC) to establish a case that stands in court.

Over 80% of survey respondents reported having faced online sexual harassment in some form, the most common being cyber stalking/being contacted repeatedly by the same person demanding a sexual relationship (44%), followed by doxxing/instances of personal information being leaked online or fake profiles being created (39%) and sexually explicit images/photos being shared with an individual without her consent (30%). Doxxing and non-consensual sharing of intimate images are not adequately addressed by the existing provisions in the IPC or the IT Act. Marshalling digital evidence in investigations into these offences is also a challenge. On their end, members of law enforcement find that their investigation is hindered by the fact that existing processes to obtain a response from platforms are cumbersome and not effective. Victims of cyber offences are thus triply victimised – by the abusers, by institutional responses of victim-blaming and by the indifference of platforms.

Against this backdrop, the ongoing law review consultation on cyber crimes against women, convened by the National Commission for Women, in December 2020 is a timely and laudable initiative especially as it coincides with two very important legislative reform processes – the complete overhaul of the ITAct (2000) proposed by the Ministry of Electronics and Information Technology and the reform of substantive and procedural aspects of criminal law spearheaded by the Ministry of Home Affairs.

2. Taking stock of existing legal remedies for gender-based hate online

Our ongoing study of legal developments to combat sexist hate speech online reveals the absence of concrete laws to combat hate speech. We have found the use of the category “hate speech” to be a recent introduction to the Indian legal discourse. In Indian law and case law, it is not found in the context of gender-based hate speech. Where sexist hate does exist, it is named by proxy through other categories of offences, such as defamation and obscenity. We found, unsurprisingly perhaps, that defamation and obscenity laws are invoked by powerful men to guard themselves (or the society they claim to represent) from the reputational harm of being associated with expressions of female sexuality that are deemed inappropriate. Our survey indicates that there is an absence of appropriate provisions that can directly respond to gender trolling and sexist abuse.

The adverse impact that gender-based cyberviolence has on victims and society at large was most recently given due recognition by the Kerala High Court in *P. Sreekumar v/s State of Kerala* in 2019, echoed in *Sreeja Prasad v/s State of Kerala* in May 2020, where the Court noted with concern the increase in intolerance and othering manifested in social media discourse.

A law against gender-based hate is only one aspect of the fight against gender-based hate speech. It is equally necessary to pay attention to procedures that are mandated for social media platforms, the regulation of misogynistic content, whether it is the notice and takedown requirements, or the turnaround time for takedown. The notice-and-takedown regime (in the existing rules as they are, as well as in the 2018 proposed amendment to the IT Act Rules of 2011) lacks a nuanced view of how liability differs according to type of content at issue (vertical differentiation) and according to the kind of function performed by the intermediary (horizontal differentiation). There is a one-size-fits-all imagination of intermediaries and their liabilities, with the result that social media intermediaries cannot be held accountable for the sharing of illegal content.

The judgment in *Shreya Singhal v/s Union of India* (2015) has hindered individual users from obtaining swift remedies to block content amounting to gendered cyber abuse. The court, given the volume of data and the number of users on various internet and social media platforms, observed that it would not be feasible for platforms to assess for harm and filter all the abusive content their platforms host. Yet, the European Commission in June 2020 observed that “On average 90% of the notifications are reviewed within 24 hours and 71% of the content is removed.” Facebook – the platform that received the largest number of notifications for hate speech – was found to have assessed notifications in less than 24 hours in 95.7% of the cases and 3.4% in less than 48 hours. A similar study of the turnaround time of large incumbent intermediaries to notifications of unlawful content under the Germany Network Enforcement Act or NetzDG demonstrated that they were able to respond to a substantial number of complaints within a 24-hour window.[1] Therefore, while crafting a legal framework to regulate such intermediaries, excessive deference to the professed feasibility issues of moderating content expressed by platform companies must be abandoned, and the burden of devising appropriate mechanisms to combat abusive content must be placed on social media intermediaries.

There is a need for legal reform in the following directions: a) legal reform that addresses sexist hate b) content governance standards grounded in privacy, dignity and the prevention of harm, c) clear

1 On the basis of transparency reporting conducted for the period January – June 2018. [C.f. Torsha Sarkar, “A Deep Dive into Content Takedown Time Frames”, The Centre for Internet and Society, November 30 2019, at pp. 10 – 11].

guidance on intermediary accountability for third party content, d) avenues for speedy redress for victims of sexist hate speech, gender-trolling and gender-based cyberviolence, e) training law enforcement to be more reliant on privacy-oriented provisions than obscenity-focused ones to provide redress against gender-based hate crimes.

3. Recommendations

3.1. Set up a committee on freedom from cyberviolence

We would like to recommend the introduction of a committee to develop a Right to Freedom from Cyberviolence Bill, 2020. This committee can also make recommendations to amend relevant provisions of the IPC, IT Act, etc., to bridge the lacuna in the laws and bring parliamentary intent on record to combat new forms of cyberviolence and amend provisions, across laws, harmoniously. The process of setting-up and operationalising such a committee must be mindful of women's dignity and privacy.

3.2. Create a law to combat gender-based hate speech online

The balance between freedom of expression and freedom from violence online needs to be re-articulated in the digital context. However, we cannot keep waiting for a new law against gender-based hate speech to be framed. As an interim measure, we recommend the introduction of the following provision in the Intermediaries Guidelines Rules to combat hate against vulnerable groups including women: "[Intermediaries] shall inform users of a computer resource not to host, display, upload, modify, publish, transmit, update or share any information that (i) spreads, incites, promotes or justifies forms of hateful expression based on religion, race, caste, place of birth, sex, sexual orientation, gender or disability. For the purposes of this provision, "hateful expression" comprises violent, dehumanising or denigrating communication.(b)(ii) promotes, endorses and incites gratuitous violence. (b)(iii) violates the privacy of individuals.(b)(iv) erodes the dignity of vulnerable groups."

3.3. Amend the Information Technology Act and other harassment laws to effectively address sexist hate online

The push back faced by the ordinance to amend the Kerala Police Act has, at the very least, demonstrated that a law to combat gender-based hate speech online needs to be brought in through a feminist, multi-constituency process. The law must have an empowered imagination of the woman

subject not only as an incidental victim in patriarchal notions of modesty, virtue and honour, but also as a person with agency, deserving of privacy and dignity.

3.3.1. Undo the unintended consequences of Shreya Singhal

The Supreme Court's judgment in *Shreya Singhal v/s Union of India* impacted two provisions of the IT Act – striking down of Section 66A on 'offensive speech' and reading down of Section 79 on 'intermediary liability'.

This has created two significant lacunae:

(a) The striking down of Section 66A has made it inordinately difficult to bring to book perpetrators of gender-trolling and sexist hate speech online. There is no equivalent provision that can counter sexist hate speech, unless it is also able to attract Section 66E or Section 67, i.e. involves sexually-explicit images or is "obscene" in nature.

(b) The reading down of "actual knowledge" in Section 79 to mean "receipt of court order" has resulted in delays in the removal of impugned content in cases of gender-based cyberviolence. For instance, when intimate images of a woman are being circulated non-consensually, the offending image/video needs to be expeditiously taken down by notifying the intermediary. If intermediaries do not respond to such user reports, there is no obvious way to compel them to do so. The *de facto* consequence of this is a gross violation of women's privacy and dignity.

3.3.2. Do away with the outdated Victorian concept of 'obscenity' used in Section 67 & 67A of the IT Act

We should discourage the application of obscenity provisions, and push for a law that instead concerns itself with gratuitously violent or non-consensual, disempowering representation of images/descriptions of women's bodies. Section 67 of the IT Act has an ambiguity of phrasing that permits it to be used against any woman who shares intimate images of herself, although it should reasonably be deployed only against non-consensual circulation of such images. It is worth noting that section 292 of the IPC, which has the same subject matter as section 67 of the IT Act, does not have a promising jurisprudence that is protective of women's dignity. Instead, section 67, with its problematic wording, has a history of being invoked in cases that deem the very representation of women's bodies or women's desire as illegitimate.[2]

2 Women in firms: Revisiting the censorship debate, Kamthan, Kumari & Pandey, Journal of Critical Reviews, Vol 6, Issue 6, 2019, http://www.jcreview.com/fulltext/197-1578480397.pdf.

3.3.3. Shift legal-juridical concepts and their interpretations to a privacy and dignity approach

Courts are demonstrating a tendency to dismiss women's right to dignity in their interactions in online spaces, deeming these as "personal", as opposed to "public" where a law against harassment may take effect. For instance, Section 294 is being interpreted as inapplicable to unwanted direct messages that are only visible to the victim of harassment. Also, in the case of *Nivrutti Hariram Gaikwad vs The State Of Maharashtra And Anr* on 11 March, 2020, where the accused sent messages to the alleged victim saying he will beat her and calling her a prostitute, the Bombay High Court concluded that "sending the personal messages on WhatsApp will not amount to utterance of obscene words in public place" and therefore that Section 294 of the IPC cannot be invoked.

Even when courts have attempted to take a victim-centred stance devoid of moral policing or shaming, the lack of epistemic categories in existing legal frameworks to address offences against privacy and dignity in online interactions proves a major impediment. Take the case of *State of West Bengal v/s Animesh Boxi @ Ani Boxi,* 2018, where the judicial magistrate in Medinipur was dealing with a case of the non-consensual circulation of intimate images. As the IPC or the IT Act in their current forms do not provide a conceptual handle to address the privacy and dignity violations stemming from such unwarranted incursion into bodily integrity, the Court was compelled to use a new, and highly unsatisfactory term, "virtual rape", which has no statutory meaning under Indian law.

3.4. Develop content governance standards grounded in privacy, dignity and the prevention of harm

We must distinguish between infringing content (content that infringes on the rights of others) and manifestly unlawful content (content that is an offence to society). Perhaps one could think of it as rights *in rem* (protections against the world) and rights *in personam* (protections against individual incursion).

This will permit us to conceive of a hierarchy for content and for consent – some content no one should be subjected to/permitted to viewing on the internet regardless of "consent" (think child pornography), some content should not be taken down just because it does not have the consent of the person involved (think whistleblowers).

3.5. Abandon the fraught legacy of “indecency” laws, towards laws grounded in privacy, autonomy and dignity

The ministry of women and child development has, at the recommendation of the National Commission for Women (NCW), proposed an amendment to the Indecent Representation of Women Act.

The current Bill proposes that indecent representation should be redefined as the depiction of the figure or form of a woman in such a way that it has the effect of being indecent or derogatory or is likely to deprave or affect public morality. It is worth noting that the Indecent Representation of Women’s Act is framed through patriarchal conceptions. There is always a threat of patriarchal morality looming on top of any framing of “decency” with regard to the representation of women’s bodies in media. As per the data available with the National Crime Records Bureau (NCRB) a total of 1898 cases have been reported under the Indecent Representation of Women (Prohibition) Act, 1986 between 2010-2014, all of which lack a common understanding of what constitutes “indecent representation”. Hence, we should not pin our hopes on the Indecent Representation of Women Act, and instead push for laws that foreground women’s privacy, autonomy and dignity.

3.6. Strengthen crime investigation without weakening encryption

Women and gender minorities have a lot to lose from poor encryption standards (hacking, stalking, doxxing, non-consensual circulation of intimate images), and from refusal of assistance from intermediaries in decryption for legitimate purposes of the state for investigation of offences committed against women. The WhatsApp traceability case is currently pending before the Supreme Court. Technological peepholes through the walls of encryption must not undermine fundamental rights to life with dignity, right to privacy and right against self-incrimination.

3.7. Set up friendlier mechanisms for reporting and reaching out to law enforcement for combating online violence

The hurdles for reporting incidents of online violence can partly be addressed by creating greater granularity in reporting mechanisms, with a wider range of options for reasons of reporting, to expand women’s ability to report harassment. This is best done in the following ways:

- Bringing in intermediary guidelines where a minimum threshold of types of content that can be “actioned” shall be clearly identified.

- A mechanism for reporting harassment, similar to the reporting mechanism in the US Digital Millennium Copyright Act (DMCA), could be set up. Here, a takedown notice claiming copyright infringing content is sent to the intermediary or Internet Service Provider, which takes action on the infringing content and informs the sharer of the takedown action. Once the content is taken down, a counter-notice can be filed by the sharer of the content. If a counter-claim is filed, the takedown is revoked after a specified time, unless the initial claimant takes legal action to keep the content taken down. Such a mechanism ensures quick action in taking down content, which is critical, especially when responding to non-consensual circulation of intimate images.
- Reporting processes should strive to be multilingual, easy to locate and understandable.
- Grievance redressal officers for cyberviolence could be specifically identified, whose office can liaison with victims and law enforcement.

3.8. Hold artificial intelligence accountable

It is important at this juncture to create accountability structures that can conduct algorithmic audits and create protocols like ensuring humans in the loop, when deploying technology-based, automated tools for proactive filtering of content. We need a new, binding global framework on social media, grounded in human rights, as we recognise efforts to build frameworks such as the five guiding principles developed by the Fairness, Accountability and Transparency in Machine Learning (FATML) community to help developers ensure algorithmic accountability. Instruments such as Santa Clara Principles and the Manila Principles that capture a growing recognition of the delimitations of intermediary responsibility, should be enacted into binding, statutory mandates. It is important to hold artificial intelligence accountable by applying a human-rights-based approach for the design, development and deployment of algorithms, and identify factors that the state and intermediaries should take into account to avoid undermining or violating human rights.[3]

3 McGregor, Lorna & Murray, Daragh & Ng, Vivian. (2019). International human rights law as a framework for algorithmic accountability. International and Comparative Law Quarterly. https://www.researchgate.net/publication/332457262