Intermediary Liability and Gender Based Violence

Report of the roundtable discussions conducted at the Digital Citizen Summit

1st November, 2018| New Delhi

By **Akriti Bopanna**
Inputs and Edited by **Ambika Tandon**

The Centre for Internet and Society, India

Introduction

Background

This report is a summary of the proceedings of the Roundtable Conference organized by the Centre for Internet and Society at the Digital Citizen Summit, an annual conference organized by the Digital Empowerment Foundation. It was conducted at the India International Centre in New Delhi on the 1st of November, 2018 from 1130AM to 1230PM.

The topic of discussion was intermediary liability and Gender Based Violence (GBV), the debate on GBV globally and in India evolving to include myriad forms of violence in online spaces in the past few years. This ranges from violence native to the digital, such as identity theft, and extensions of traditional forms of violence, such as online harassment, cyberbullying, and cyberstalking[1]. Given the extent of personal data available online, cyber attacks have led to a variety of financial and personal harms.[2] Studies have explored the extent of psychological and even physical harm to victims, which has been found to have similar effects to violence in the physical world[3]. Despite this, technologically-facilitated violence is often ignored or trivialised. When present, redressal mechanisms are often inadequate, further exacerbating the effects of violence on victims.

TheRoundtable explored ways of how intermediaries can help tackle gender based violence and discussed attempts at making the Internet a safer place for women which can ultimately help make it a gender equal environment. It also analyzed the key concerns of privacy and security leading the conversation to how we can demand more from platforms for our protection and how best to regulate them.

The roundtable had four female and one male participants from various civil society organisations working on rights in the digital space.

---

[1] *See* Khalil Goga, "How to tackle gender-based violence online", World Economic Forum, 18 February 2015, <https://www.weforum.org/agenda/2015/02/how-to-tackle-gender-based-violence-online/>. *See also* Shiromi Pinto, "What is online violence and abuse against women?", 20 November 2017, Amnest International, <https://www.amnesty.org/en/latest/campaigns/2017/11/what-is-online-violence-and-abuse-against-women/>.

[2] Nidhi Tandon, et. al., "Cyber Violence Against Women and Girls: A worldwide wake up call", UN Broadband Commission for Digital Development Working Group on Broadband and Gender, <http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/images/wsis/GenderReport2015FINAL.pdf>

[3] *See* Azmina Dhrodia, "Unsocial Media: The Real Toll of Online Abuse against Women", Amnesty Global Insights Blog, <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4>

Roundtable discussion

a) Online abuse

The discussion commenced with the acknowledgement of it being well documented that women and sexual minorities face a disproportionate level of violence in the digital space, as an extension/reproduction of physical space. GBV exists on a continuum from the physical, verbal, and technologically enabled, either partially or fully, with overflowing boundaries and deep interconnections between different kinds of violence. Some forms of traditional violence such as harassment, stalking, bullying, sex trafficking, extend themselves into the digital realm while other forms are uniquely tech enabled like doxxing and morphing of imagery. Due to this considerations of anonymity, privacy, and consent, need to be re-thought in the context of tech enabled GBV. These come into play in a situation where the technological realm has largely been corporatised and functions under the imperative of treating the user and their data as the final product.

It was noted early on that GBV online can be a misnomer because it can be across a number of spaces and, the participants concentrated on laying down the specific contours of tech mediated or tech enabled violence. One of the discussants stated that the term GBV is a not a useful one since it does not encompass everything that is talked about when referring to online abuse. The phenomenon that gets the most traction is trolling on social media or abuse on social media. This is partly because it is the most visible people who are affected by it, and also since often, it is the most difficult to treat under law. In a 2012 study by the Internet Democracy Project focusing on online verbal abuse in social media, every woman they interviewed started by asserting that she is not a victim. The challenge with using the GBV framework is that it positions the woman as a victim. Other incidents on social media such as verbal abuse where there are rape threats or death threats, especially when there is an indication that the perpetrator is aware of the physical location of the victim, need to be treated differently from say online trolling.

Further, certain forms of violence, such as occurrences of ‘revenge porn’ or the non-consensual sharing of intimate images, including rape videos are easier to fit within the description of GBV. It is important to make these distinctions because the remedies then should be commensurate with perceived harm. It is not appropriate to club all of these together since the criminal threshold for each act is different. Whereas being called a “slut” or a “bitch” would not be enough for someone to be arrested, if a woman is called that repetitively by a large number of people the commensurate harm could be quite significant. Thus, using GBV as a broad term for all forms of violence ends up invisiblising certain forms of violence and prevents a more nuanced treatment of the discussion.

In response to this, a participant highlighted the normalisation of gendered hate speech, to the extent of lack of recognition as a form of hate speech. This lacunae in our law stems from the fact that we inherited our hate speech laws from a colonial era where it was based on the grounds of incitement of violence, more so physical violence. As a result, we do not take the International Covenant on Civil and Political Rights (ICCPR) standard of incitement to discrimination. If the law was based on an incitement to discriminate point of view then acts of trolling could come under hate speech. Even in the United Kingdom where there is higher sentencing for gender based crime as compared to other markers of identity such as race, gender does not fall under the parameters of hate speech. This can also be attributed to the threshold at which criminalization kicks in for such acts.

A significant aspect of online verbal abuse pointed out by a participant was that it does not affect all women equally. In a study, the Twitter accounts of 12 publicly visible women across the political spectrum were looked at for 2 weeks in early December, 2017. They were filtered against keywords and analyzed for abusive content. One Muslim woman in the study had extremely high levels of abuse, being consistently addressed as “Jihad man, Jihad didi or Jihad biwi”. According to the participant, she is also the least likely to get justice through the criminal system for such vitriol and as such, this disparity in the likelihood of facing online abuse and accessing official redressal mechanisms should be recognized. Another discussant reaffirmed the importance of making a distinction between online abuse against someone as opposed to gender based violence online where the threat itself is gendered.

In a small ethnographic study with the Bangalore police undertaken by one of the participants, the police were asked for their opinion on the following situation: A women voluntarily providers photos of herself in a relationship and once the relationship is over, the man distributes it. Is there a cause for redressal?
Policemen responded that since she gave it voluntarily in the first instance, the burden of the consequences is now on her. So even in a feminist framework of consent and agency where we have laws for actions of voyeurism and publishing photos of private parts, it is not being recognized by institutional response mechanisms.

b) Intermediary Liability

Private communications based intermediaries can be understood to be of two types: those that enable the carriage/transmission of communications and provide access to the internet, and those that host third party content. The latter have emerged as platforms that are central to the exercising of voice, the exchange of information and knowledge, and even the mobilisation of social movements. The norms and regulations around what constitutes gender based violence in this realm is then shaped not only by state regulations, but content moderation standards of these intermediaries. Further, the kinds of preventive tools and tools providing redressal are controlled by these platforms. More than before, we are looking

deeper into the role of these companies that function as intermediaries and control access to third party content without performing editorial functions.

In the Intermediary Liability framework in the United States formulated in the 1990s, the intermediaries that were envisioned were not the intermediaries we have now. With time, the intermediary today is able to access and possess your data while urging a certain kind of behaviour from you. There is then an intermediary design duty which is not currently accounted for by the law. Moreover, the law practices a one size fits all regime whereas what could be more suitable is having approached tailored as per the offence. So for child pornography, a 'removal when uploaded' action using artificial intelligence or machine learning is appropriate but a notice and takedown approach is better for other kinds of content takedown.

Globally, another facet is that of safe harbour provisions for platforms. When intermediaries such as Google and Facebook were established, they were thought of as neutral pipes since they were not creating the content but only facilitating access. However, as they have scaled and as their role in ecosystem has increased, they are now one of the intervention points for governments as gatekeepers of free speech. One needs to be careful in asking for an expansion of the role and responsibilities of platforms because then complementary to that we will also have to see that the frameworks regulating them need to be revisited. Additionally, would a similar standard be applicable to larger and smaller intermediaries, or do we need layers of distinction between their responsibilities? Internet platforms such as the GAFA (Google, Apple, Facebook and Amazon) yield exceptional power to dictate what discourse takes place and this translates into the the online and offline divide disappearing. Do we then hold these four intermediaries to a separate and higher standard? If not, then all small players will be held to stringent rules disadvantaging their functioning and ultimately, stifling innovation. Thus, regulation is definitely needed but instead of a uniform one, one that's layered and tailor-made to different situations and platform visibility levels could be more useful.

Some participants shared the opinion that because these intermediaries are based in foreign countries and have primary legal obligations there, the insulation plays out in the citizen's benefit. It lends itself a layer of freedom of speech and expression that is not present in the substantive law, rule of law framework or the institutional culture in India.

Child pornography is an area where platforms are taking a lot of responsibility. Google has spoken about how they have been using machine learning algorithms to block 40% of such content and Microsoft is also working on a similar process. If we argue for more intervention from platforms, we simultaneously also need to look at their machine learning algorithms. Concerns of how these algorithms are being deployed and further, being incorporated into the framework of controlling child pornography are relevant since there is not much accountability and transparency regarding the same.

Another fraction that has emerged from recent events is the divide between traditional form of media and new media. Taking the example of rape victims and sexual harassment claims, there are strict rules regarding the kinds of details that can be disclosed and the manner in which this is to be done. In the Kathua rape case, for instance, the Delhi High Court sent a notice to Twitter and Facebook for revealing details because there are norms around this even though they have not been applicable to platforms. Hence, there are certain regulations that apply to old media that have now escaped in the frameworks applicable to the new media and at some level that gap needs to be bridged.

c) Role of Law

One of the participants brought up the question; what is the proper role of the law and does it come first or last? In case of the latter, the burden then falls upon the kind of standard setting that we do as a society. The role of platforms as an entity in mediating the online environment was discussed, given the concerns that have been highlighted about this environment, especially for women. The third thing to be considered is whether we run the risk of enforcing patriarchal behaviour by doubling down on the either of the two aforementioned factors. If legal standards are made too harsh they may end up reinforcing a power structure that is essentially dominated by upper caste men who comprise a majority of staff within law enforcement and the judiciary. Even though the subordinate judiciary do have mahila courts now, the application of the law seems to reify the position of the woman as the victim. This also brings up the question of who can become a victim within such frameworks, where selective bias such as elements of chastity come to play as court functions are undertaken.

An assessment of the way criminal law in India is used to stifle free speech was carried out in 2013 and repeated in 2018, illustrating how censorship law is used to stifle voices of minorities and people critical of the political establishment. Even though it is perhaps time to revisit the earlier conceptualizations of intermediaries as neutral pipes, it is concerning to look at the the court cases regarding safe harbour in India. Many of them are carried out with the ostensible objective of protecting women's rights. In *Kamlesh Vaswani V Union of India*, the petition claims that porn is a threat to Indian women and culture, ignoring the reality that many women watch porn as well. Pornhub releases figures on viewership every year, and of the entirety of Indian subscribers one third are women. This is not taken into account in such petitions. In *Prajwala V Union of India,* an NGO sent the Supreme Court a letter raising concerns about videos of sexual violence being distributed on the internet. The letter sought to bring attention to the existence of such videos, as well as their rampant circulation on online platforms. At some point in the proceedings, the Court wanted the intermediaries to use keywords to take down content and keeping aside poor implementation, the rationale behind such a move is problematic in itself. For instance, if you choose sex as one of those words then all sexual education will disappear from the Internet. There are many problems with court encouraged filtering systems like one where a system automatically tells you when

a rape video goes up. The question arises of how will you distinguish between a video that was consensually made depicting sexual activities and a rape video. The narrow minded responses to the Sabu Mathew and Prajwala cases originate in the conservative culture regarding sexual activity prevalent in India.

In a research project undertaken by one of the participants in the course of their work, they made a suggestion to include gender, sexuality and disability as grounds for hate speech while working with women's rights activists and civil society organisations. This suggestion was not well received as they vehemently opposed more regulation. In their opinion, the laws that India has in place are not being upheld and creating new laws will not change if the implementation of legislation is flawed. For instance, even though the Supreme Court stuck down S.66A, Internet Freedom Foundation has earlier provided instances of its continued usage by police officers to file complaints.[4] Hate speech laws can be used to both ends, even though unlike in the US they do not determine whose speech they want to protect. Consequently, in the US a white supremacist gets as much protection as a Black Lives Matter activist but in India, that is not the case. The latest Law Commission Report on hate speech in India tries to make progress by incorporating the ICCPR view of incitement to discriminate and include dignity in the harms. It specifically speaks about hate speech against women saying that it does not always end up in violence but does result in a harm to dignity and standing in society. Often, protectionist forms of speech such as hate speech often end up hurting the people it aims to protect by reinforcing stereotypes.

Point of View undertook a study where they looked at the use of S.67 in the Information Technology (IT) Act which criminalizes obscene speech when you use a medium covered by the IT, in which they found that the section was used to criminalize political speech. In many censorship cases, the people who those provisions benefit are the ones in power.[5] For instance in S.67, obscenity provisions do not protect women's rights, they protect morality of society. Even though these are done in the name of protecting women, when a woman herself decides that she wants to publish a revealing picture of herself online, it is disallowed by the law. That kind of control of sexuality is part of a larger patriarchal framework which does not support women's rights or recognise her sexuality. However, under Indian law, there are quite a few robust provisions for image based abuse, and there is some recognition of women in particular being vulnerable to it. S.66A of the IT Act specifically recognizes that it is a criminal activity to share images of someone's private parts without their consent. This then also encompasses instances of 'revenge porn'. That provision has been in place in India since 2008, in contrast to the US where half the states still do not have such a provision. Certain kinds of vulnerability have adequate recognition in the law, thus one should be wary of calls of censorship and lowering the standards for criminalizing speech.

---

[4] *See* Abhinav Sekhri and Apar Gupta, "Section 66A and other legal zombies", Internet Freedom Foundation Blog, <https://internetfreedom.in/66a-zombie/?

[5] See Bishakha Datta "Guavas and Genitals", Point of View <https://itforchange.net/e-vaw/wp-content/uploads/2018/01/Smita_Vanniyar.pdf>

d) Non-legal interventions

This section centres around the discussions of redressal mechanisms that can be used to address some of the forms of violence which do not emanate from the law. All of the participants emphasized the importance of creating safe spaces through non-legal interventions. It was debated whether there is a need to always approach the law or if it is possible to categorize forms of online violence according to the gravity of the violation committed. These can be in the form of community solutions where law is treated as the last resort. For instance, there was support for using community tools such as 'feminist trollback' where humor can be used to troll the trolls. Trolls feed on the fear of being trolled, so the harm can be mitigated by using community initiatives wherein the target can respond to the trolls with the help of other people in the community. It was reiterated that non technical and legal interventions are needed not only from the perspective of power relations within these spaces but also access to the spaces in the first place. Accordingly, the government should work on initiatives that get more women online and focus on policies that makes smartphones and data services more accessible. This would also be a good method to increase the safety of women and benefit from the strength in numbers.

In cases of the non-consensual sharing of intimate images, law can be the primary forum but in cases of trolling and other social media abuse, the question was raised - should we enhance the role of the intermediary platforms? Being the first point of intervention, their responsibility should be more than it currently is. However this would require them to act in the nature of police or judiciary and necessitate an examination of their algorithms. A large proportion of the designers of such algorithms are white males, which increases the possibility of their biases against women of colour for instance, to feed into the algorithms and reinforce a power structure that lacks accountability.

Participants questioned the lack of privacy in design with the example in mind being of how registrars do not make domain owner details private by default. Users have to pay an additional fee for not exposing their details to public and the notion of having to pay for privacy is unsettling. There is no information being provided during the purchasing of the domain name about the privacy feature as well. It was acknowledged that for audit and law enforcement purposes it is imperative to have the information of the owner of a domain name and their details since in cases of websites selling fake medicines, arms or hosting child pornography. Thus, it boils down to the kind of information necessary for law enforcement. Global domain name rules also impact privacy on the national level. The process of ascertaining the suitability and necessity of different kinds of information excludes ordinary citizens since all the consultations take place between the regulatory authority and the state. This makes it difficult for citizens to participate and contribute to this space without government approval.

Issues were flagged against community standards in that the violence that occurs to women is also because the harms are not equal for all. Further, some users are targeted specifically because of the community they come from or the views they have. Often also because, they represent a ‘type’ of a woman that does not adhere to the ‘ideal’ of a woman held by the perpetrator. Unfortunately community standards do not recognise differential harms towards certain communities in India or globally. Twitter, for example, regularly engages in shadow banning and targets people who do not conform to the moral views prevalent in that society where the platform is engaging in censorship. We know these instances occur only when our community members notice and notify us of the same. There is a certain amount of labor that the community has already put in flagging instances of these violations to the intermediary which also needs recognition. In this situation, Twitter is disproportionately handling how it engages with the two entities in question. Community standards could thus become a double edged sword without adding additional protections for certain disadvantaged communities.

Conclusion

Currently, intermediaries are considered neutral pipes through which content flows and hence have no liability as long as they do not perform editorial functions. This has also been useful in ensuring that the freedom of speech is not harmed. However, given their potential ability to remedy this problem, as well as the fact that intermediaries sometimes benefit financially from such activities, it is important to look at the intermediaries' responsibility in addressing these instances of violence. Governments across the world have taken different approaches to this question[6]. Models, such as in the US, where intermediaries have been solely responsible to institute redressal mechanisms have proven to be ineffectual. On the other hand, in Thailand, where intermediaries are held primarily liable for content, the monitoring of content has led to several free speech harms.

People are increasingly looking at other forms of social intervention to combat online abuse since technological and legal ones do not completely address and resolve the myriad issues emanating from this umbrella term. There is also a need to make the law gender sensitive as well as improving the execution of laws at ground level, possibly through sensitisation of law enforcement authorities. Gender based violence as a catchall phrase does not do justice to the full spectrum of experiences that victims face, especially women and sexual minorities. Often these do not attract criminal punishment given the restricted framework of the current law and need to be seen through the prism of hate speech to strengthen these provisions.

Some actions within GBV receive more attention than others and as a consequence, these are the ones platforms and governments are most concerned with regulating. Considerations of free speech and censorship and the role of intermediaries in being the flag bearers of either has translated into growing calls for greater responsibility to be taken by these players. The roundtable raised some key concerns regarding revisiting intermediary liability within the context of the scale of the platforms, their content moderation policies and machine learning algorithms.

[6] 'Examining Technology-Mediated Violence Against Women Through a Feminist Framework: Towards appropriate legal-institutional responses in India', Gurumurthy et al., January 2018.