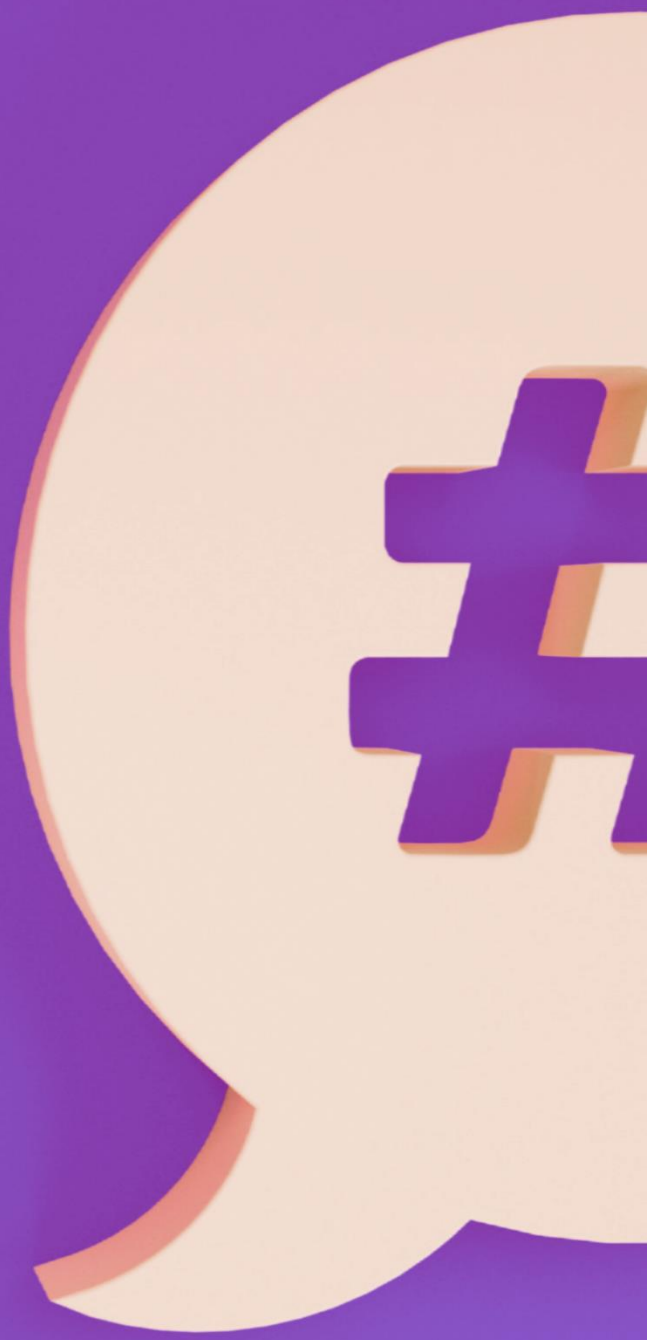
Rethinking Legal-Institutional Approaches to Sexist Hate Speech in India

Towards Holistic Regulation of Online Hate Speech

Rishab Bailey and
Vrinda Bhandari



IT for Change
February 2021

Towards Holistic Regulation of Online Hate Speech

Rishab Bailey and Vrinda Bhandari

Rishab Bailey is a researcher at the National Institute of Public Finance and Policy, Delhi.

Vrinda Bhandari is an Advocate at the Supreme Court of India and Delhi High Court.

This paper is part of a series under IT for Change's project, Recognize, Resist, Remedy: Combating Sexist Hate Speech Online. The series, titled Rethinking Legal-Institutional Approaches to Sexist Hate Speech in India, aims to create a space for civil society actors to proactively engage in the remaking of online governance, bringing together inputs from legal scholars, practitioners, and activists. The papers reflect upon the issue of online sexism and misogyny, proposing recommendations for appropriate legal-institutional responses. The series is funded by EdelGive Foundation, India and International Development Research Centre, Canada.

February, 2021

Conceptualisation

Anita Gurumurthy, Nandini Chami, Bhavna Jha

Editors

Anita Gurumurthy, Bhavna Jha

Editorial Support

Amay Korjan, Ankita Aggarwal, Sneha Bhagwat, Tanvi Kanchan

Design and Layout

Sneha Bhagwat

Rishab Bailey and Vrinda Bhandari

Towards Holistic Regulation of Online Hate Speech

Introduction

In January 2020, Amnesty International India released a report reviewing the content of thousands of tweets sent to 95 female Indian politicians on Twitter and found that they were subject to unprecedented levels of online abuse, especially if they were outspoken and forthright. The women were targeted not only for their views expressed online, but also based on their identities such as gender, religion, caste, and marital status.[1] Online abuse, however, is not restricted to women politicians. A brief look at the Twitter timelines of journalists such as Rana Ayyub or Barkha Dutta, the news coverage surrounding Rhea Chakraborty, or the rape threats made to M.S. Dhoni's six-year-old daughter over his poor IPL performance[2] reveals the misogyny that is ingrained in our society. Anonymity (or the sense of relative obscurity), a perceived majority status, and increased desensitisation, fuels more toxic behaviour online, exacerbating the patriarchy and misogyny in the offline world, in part because most actions online lack social consequences.[3]

India, as with other jurisdictions, has been struggling to deal with the issue of how to regulate speech in the online ecosystem. Regulatory responses have usually seen demands for either new criminal laws to be implemented or for intermediaries to 'do more' to make the online ecosystem safer.[4] The

[1] Amnesty International India, "Shocking scale of abuse on Twitter against women politicians in India", January 2020, https://www.amnestyusa.org/press-releases/shocking-scale-of-abuse-on-twitter-against-women-politicians-in-india/

[2] Web Desk, "MS Dhoni's 5-year-old daughter Ziva gets rape threats after CSK loses match to KKR", Deccan Herald, 11th October 2020, https://tinyurl.com/yxebaujk

[3] See generally, Jesse Fox, Why the online trolls troll, Psychology Today, 12 August 2014, https://www.psychologytoday.com/us/blog/better-living-technology/201408/why-the-online-trolls-troll

[4] Bahl, Bailey and Rahman, "Internet intermediaries and online harms", Working Paper 06, Data Governance Network, March 2020, https://www.datagovernance.org/files/research/BahlRahmanBailey_-_Paper_6-2.pdf

latter is often predicated on obligations cast under the safe harbour framework in Section 79 of the Information Technology (IT) Act, 2000.[5]

In order to address online hate speech,[6] the legal framework in India broadly requires:

a. Individuals to file complaints before State authorities under relevant criminal laws, who can then exercise powers under the IT Act or the Indian Penal Code (IPC) to seek blocking of content and/or prosecution of those perpetrating illegal acts.

b. Intermediaries to warn users against committing various illegal acts while using their services,[7] to assist law enforcement,[8] and to put in place grievance redress mechanisms to address user complaints within a period of one month.[9]

c. Certain intermediaries (Internet Service Providers and search engines) to filter content concerning rape, paedophilia or pre-natal sex determination, an obligation that has been imposed through judicial decisions.[10]

Importantly, in the first case above, the onus is largely on the victim to take forward a process that, more often than not, fails to function in a reasonably efficacious manner (let alone with due regard to gender related issues). The criminal justice system in India is under-resourced, under-skilled, encoded with patriarchal notions of morality, and suffers from systemic shortcomings. For most women, filing police complaints pertaining to online hate speech is time-consuming, harassing, expensive, and ultimately does not result in police action and successful prosecution. Moreover, the interface between the criminal justice system and intermediaries continues to be poor.[11] The ad hoc and often

[5] Section 79 of the Information Technology Act, 2000, absolves intermediaries for illegal acts committed on their platforms if they do not contribute thereto, take remedial action when required through lawful process, and follow due diligence guidelines prescribed by the Government. This provision is substantially different from the "safe harbour" provisions in American law (the Communications Decency Act, 1996), which specifically establishes a self-regulatory system for online content moderation. Indian law, as interpreted in the Shreya Singhal case, instead focuses on State-led censorship processes. Bailey, Parsheera, Rahman, "Comments on the (draft) Information Technology (Intermediaries Guidelines Amendment) Rules, 2011, January 31, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3328401

[6] We use the phrase "hate speech" very broadly in this piece to refer generally to all online misogynistic, hateful and abusive speech. We do not restrict the phrase only to instances where the speech may produce imminent illegal action or otherwise provoke violence, but include the panoply of discriminatory and harmful speech that arises out of patriarchal social systems.

[7] Rule 3, Information Technology (Intermediaries Guidelines) Rules, 2011, (hereinafter the "Intermediary Guidelines") requires intermediaries to proscribe sharing of content that is grossly harmful, harassing, defamatory, obscene, pornographic, paedophilic, invasive of privacy, or otherwise unlawful.

[8] By providing any information requested through a lawful process, including verifying user identities, or providing any other assistance required to prevent or investigate/prosecute offences. Refer Rule 7, Intermediary Guidelines.

[9] A failure to adhere to these conditions could see the intermediary concerned be held responsible for the illegal act. See Section 79, IT Act, 2000.

[10] *In Re: Prajwala*, SMW Cr. No. 3/2015, Supreme Court of India; *Sabu Mathew George v. Union of India*, (2017) 2 SCC 514; *Kamlesh Vaswani v. Union of India*, 2019 (1) CTC 548.

[11] Intermediaries are often accused of not promptly assisting law enforcement. They in turn point to problems such as the lack of due process followed by State agencies, the often ambiguous or over-broad nature of requests for assistance, etc. See Smriti Parsheera and Prateek Jha, "Cross-border data access for law enforcement: What are India's strategic options?", Carnegie India Working Paper (forthcoming).

highly political nature of interventions by legislative bodies is not conducive to consistent regulatory outcomes.[12] Approaching platforms for redress can also be challenging, as grievance redress mechanisms can be unresponsive and lack transparency. A significant problem is the lack of consistent application of a platform's terms and conditions.[13] Platforms may also benefit monetarily from promoting hateful speech, thereby creating a perverse incentive to avoid self-regulation.[14]

> For most women, filing police complaints pertaining to online hate speech is time-consuming, harassing, expensive, and ultimately does not result in police action and successful prosecution. Importantly, the interface between the criminal justice system and intermediaries continues to be poor.

In this context, this paper provides an overview of regulatory interventions that have been proposed to deal with the issue of online hate speech in India, and argues that attempts to find a 'silver bullet' solution are unlikely to bear fruit. The focus on either creating new criminal offences or casting substantive policing obligations on intermediaries may lead to inequitable results and unintended consequences. These could negatively affect the very groups these interventions aim to protect. The issue of regulating online hate speech is complex, and therefore we suggest that regulatory interventions must (a) take into account the range of modalities available and (b) focus on process improvements.

[12] See for example, IANS, "Parliamentary panel hints at action after Twitter refuses to attend hearing", NDTV, February 11, 2019, https://tinyurl.com/y6mtwymo; The Federal, "Facebook gives Delhi panel the slip, refuses to attend meet", September 15, 2020, https://thefederal.com/news/facebook-gives-delhi-panel-the-slip-refuses-to-attend-meet/; MK Venu and Maya Mirchandani, "Political mudslinging over Facebook bias will only let the social media giant off the hook", The Wire, September 7, 2020, https://thewire.in/tech/facebook-political-bias-mark-zuckerberg

[13] The terms and conditions of most large platforms do indeed proscribe hate speech. The problem, however, as noted by the Madras High Court while temporarily banning the use of TikTok, is the unwillingness of platforms to apply these on a consistent basis. See order of the Madras High Court of April 3, 2019, in *S Muthukumar v Telecom Regulatory Authority of India*, WP (MD) No. 7855/2019. Further, complaints processes can often be difficult to manoeuvre (for example, in terms of trying to fit a complaint into pre-scripted formats that may require identification of a specific offence), particularly for those without adequate digital literacy, and can also lack transparency. Users will often not be informed of why any particular action was taken (or not taken), etc.

[14] Jon Evans, "Facebook is broken", Tech Crunch, June 4, 2017, https://tinyurl.com/y827xh3q; Jon Evans, "Facebook isn't free speech, its algorithmic amplification optimized for outrage", October 20, 2019, https://tinyurl.com/yyltbs92; Peter Dzikes, "Study: On Twitter, false news travels faster than true stories", MIT News Office, March 8, 2018, https://tinyurl.com/y57v78vn; Safiya Umoja Noble, "Algorithms of Oppression: How search engines reinforce racism", NYU Press, 2018. Also see Newley Purnell and Jeff Horowitz, "Facebook's Hate Speech Rules Collide with India's Politics", The Wall Street Journal, August 14, 2020, https://tinyurl.com/yxtqgm2h

Initiatives to Regulate Intermediaries

As is evident from the rising instances of online hate speech, the methods detailed above have largely proved inadequate.[15] Accordingly, a number of proposals have been made in recent years – in the form of legal or policy instruments,[16] interventions by courts,[17] and recommendations of government committees.[18] These focus on introducing new laws and processes for online platforms to follow, notably:

a. That intermediaries should implement automated filtering mechanisms.

b. That intermediaries should create backdoors to encryption or implement traceability mechanisms or identity verification systems.

c. That third parties should be empowered to monitor the internet and assist with identification of illegal content.

d. That the State should increase its efforts to block illegal content or even entire platforms, including through the establishment of independent censorship mechanisms.

However, each suffers from the same general problem in that they focus on attempting to find 'silver bullet' solutions, thereby inevitably prescribing the use of blunt instruments to tackle what are complex problems. Further, the suggestions are usually paternalistic in nature. Invariably, they lack a thorough consideration of interests (not least due to the frequent lack of stakeholder consultation), and tend to prescribe methods that could create significant unintended effects. Requirements to implement (AI-based) filtering mechanisms for instance, can raise questions arising from privatising and automating censorship functions. We also have to consider issues concerning algorithmic accountability and possible discriminatory effects thereof, as well as the costs for smaller players in

[15] The National Crime Records Bureau (NCRB) reports 4242 cases of cybercrimes against women in 2017, 6030 in 2018, and 8379 in 2019. See NCRB, "Table 9A.10, Cyber Crimes Against Women (State/UT Wise)", Crime in India: Table Contents (from 2017-2019), Government of India, https://ncrb.gov.in/en/crime-in-india-table-addtional-table-and-chapter-contents. Further, a number of academic and anecdotal sources point to increasing online misogyny, including as a result of the increasing number of women who now have access to the internet as well as increasing politicisation of online forums. See Sandra Handy, "The rise of misogynistic online forums and their increasing politicisation", The Science Times, April 20, 2019, https://tinyurl.com/y32g4urm, Debbie Ging and Eugenia Siapera, "Special issue on online misogyny", Feminist Media Studies, 18(4), 515-524, https://tinyurl.com/y5673zr6

[16] Notable examples include the publication of the draft Information Technology (Intermediaries Guidelines Amendment) Rules, 2018, and a private members' bill known as the Social Media Accountability Bill, 2018.

[17] Notably in cases such as *Sabu Mathew*, *In Re: Prajwala*, *Kamlesh Vaswani* (cited previously) and *Anthony Clement Rubin v. Union of India*, WP No 20774 and 20214/2018, Madras High Court.

[18] For example, the ad-hoc Committee of the Rajya Sabha, 2020, Expert Committee of Secretaries, 2018, and the TK Vishwanathan Committee, 2017. See generally, Bahl, Bailey and Rahman, supra note 4.

the ecosystem.[19] The broad nature of the suggested interventions also leaves them open to misuse, or in fact, to harm the very groups they are seeking to protect. For instance, the creation of civil-society based censorship organisations is not necessarily advisable in a country where patriarchal moral vigilantism is not uncommon.[20] The recent suggestions of an ad hoc committee of the Rajya Sabha,[21] to permit backdoors in encryption, so as to enable tracing of paedophilic content, is similarly questionable in that it could lead to *lesser* security for the very groups the policy is seeking to protect.[22] Implementing identification requirements can create barriers to access for women and other vulnerable groups in society, while also taking away from the right to anonymity (which can be an important weapon of empowerment).[23]

Problems with Criminalisation

The focus on criminalising online hate speech ignores the misuse that is possible due to the broad/vague wording of provisions that make them susceptible to arbitrary application, usually to re-entrench the prejudices of the existing patriarchal system. Section 66A of the IT Act was (in)famously justified by the government as necessary to fight harassment and online abuse, and ended up being used to suppress legitimate and often political speech.[24]

---

[19] Susan Leavy, "Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning", IEEE/ACM 1st International Workshop on Gender Equality in Software Engineering (GE), Gothenburg, 2018, pp. 14-16, https://ieeexplore.ieee.org/document/8452744, Josh Feast, "Four ways to address gender bias in AI", Harvard Business Review, November 20, 2019, https://hbr.org/2019/11/4-ways-to-address-gender-bias-in-ai, Smriti Parsheera, "Adoption and regulation of facial recognition technologies in India: Why and why not?", Data Governance Network, Working Paper 05, November 2019, https://datagovernance.org/files/research/NIPFP_Smriti_FRT_-_Paper_5.pdf, Surya Deva, "Addressing the gender bias in artificial intelligence and automation", OpenGlobalRights, April 10, 2020, https://tinyurl.com/y6ayl8n3. Also see Bailey, Parsheera, Rahman, supra note 5.

[20] The government however has recently established a system for individuals to sign up as "cyber warriors" to help identify and report illegal content. See "Cyber Crime Volunteers Concept", Ministry of Home Affairs, Government of India, last accessed on October 11, 2020, https://cybercrime.gov.in/Webform/cyber_volunteers_concept.aspx

[21] "Rajya Sabha Committee calls for mandatory apps on all devices and filters to regulate children's access to pornography content", Press Information Bureau, Government of India, January 25, 2020, https://pib.gov.in/PressReleseDetail.aspx?PRID=1600505

[22] Bailey, Bhandari and Rahman, "Backdoors to encryption: Analysing an intermediary's duty to provide technical assistance", Data Governance Network (forthcoming); Wofgang Schulz and Joris van Hoboken, "Human rights and encryption", UNESCO, 2016, https://unesdoc.unesco.org/ark:/48223/pf0000246527; ISOC, "Encryption: Essential for the LGBTQ+ community", November 1, 2019, https://tinyurl.com/y2cuold2; EDRi, "Women's rights online: Tips for a safer digital life", March 8, 2019, https://edri.org/our-work/womens-rights-online-tips-for-a-safer-digital-life/; United Nations Human Rights Council, "Written statement submitted by Association for Progressive Communications", A/HRC/36/NGO/88, August 31, 2017, https://www.apc.org/sites/default/files/G1725558_0.pdf

[23] "On the internet, the option of anonymity enables women to construct their identity on their own terms. They can discuss issues that are otherwise considered too 'sensitive' for the public domain...For many, the Internet is a "safe space" to vent against a repressive gender regime in the offline world." Shruti Jain, "The rising fourth wave: Feminist activism on digital platforms in India", Observer Research Foundation, ORF Issue Brief No. 384, July 2020, https://tinyurl.com/yydfwt2c. While anonymity has and does create problems (such as that of assigning responsibility for illegal actions), putting in place broad based identity verification requirements is arguably, a disproportionate response. Bailey, Bhandari and Rahman, "Traceability, identification and anonymity on the Internet", Data Governance Network (forthcoming).

[24] Richa Kaul Padte, "Section 66A, sexual harassment, and women's rights", Internet Democracy Project, December 2012, https://internetdemocracy.in/2012/12/section-66a-sexual-harassment-and-womens-rights-2/

Where does that leave us? Do we need separate laws to deal with abuse online? The State of Chhattisgarh, answering the question in the affirmative, recently introduced a State amendment to the IPC, by inserting Section 509B to the IPC, criminalising sexual harassment online. On the face of it, this may seem like a laudatory step, but a closer analysis leaves a cause for concern. The section criminalises the making of "obscene, lewd, lascivious, filthy or indecent" comments online with an intent to harass or cause annoyance or mental agony to the victim.[25] The use of the terms "obscene", "lewd", "lascivious" are not new in Indian law, and build upon existing laws that criminalise obscenity and sexually explicit acts under Sections 67 and 67A of the IT Act and Sections 292-294 of the IPC.[26] These laws have been criticised as reflecting outmoded Victorian notions of public morality that punish deviations from stereotypical traditional notions of public morality and 'Indian culture'.[27] Such regulation of sexual speech is predicated on the assumption that "any representation of sex and sexuality is bad; sexually explicit representations are immoral; and sex and sexuality are not a normal part of our humanity, but a corrupting and unhealthy influence from which 'decent people' must be protected".[28]

The problematic underpinning of our laws governing obscenity is accompanied by a complete side-stepping of the issue of consent, and unchecked misuse, that has resulted in increased censorship of sexual expression and experiencing sexual pleasure online, while reducing voluntary sexual expression and freedom of speech online.[29]

> The problematic underpinning of our laws governing obscenity is accompanied by a complete side-stepping of the issue of consent, and unchecked misuse that has resulted in increased censorship of sexual expression and experiencing sexual pleasure online while reducing voluntary sexual expression and freedom of speech online.

Section 509B also uses the terms "filthy" or "indecent" that are undefined, vague, and open to subjective (mis)interpretation. The misuse of these laws was evident in the recent case involving

[25] Section 509B of the IPC, as inserted by the Criminal Law (Chhattisgarh Amendment) Act, 2013 http://www.bareactslive.com/Ch/cg240.htm

[26] Section 67 of the IT Act criminalises the publication or transmission of obscene material electronically, while Section 67A prohibits the publication or transmission of material containing sexually explicit acts electronically. Sections 292 to 294, IPC punish the sale, etc. of obscene books and song.

[27] Gautam Bhatia, "Offend, shock or disturb: Free speech under the Indian constitution", (Oxford University Press, 2016); Richa Kaul Padte and Anja Kovacs, "Keeping women safe? Gender, online harassment and Indian law", Internet Democracy Project, 2013, https://internetdemocracy.in/reports/keeping-women-safe-gender-online-harassment-and-indian-law/

[28] Ratna Kapur, "Who draws the line? Feminist reflections on speech and censorship", 31(16/17) Econ. and Pol. Weekly WS 15 (1996).

[29] See generally, Bishaka Dutta, "Guavas and genitals: A research study in Section 67 of the Information Technology Act, 2018", https://itforchange.net/e-vaw/wp-content/uploads/2018/01/Smita_Vanniyar.pdf

Mohammed Zubair, a journalist and co-founder of the popular fact-checking online news outlet, ALT News.

Given the political nature of his work, Zubair has often been the target of abuse on social media, and was particularly attacked by one Jagdish Singh.[30] Singh's Twitter profile display picture was a photo of him with his minor granddaughter. In response to Singh's abusive tweet in August 2020, Zubair re-tweeted the tweet to Singh's wall, and after blurring/pixelating the face of the minor girl, referenced Singh's display picture and commented, "Hello Jagdish Singh, Does your cute granddaughter know about your part time job of abusing people on social media? I suggest you to change your profile pic."

Jagdish Singh immediately filed a complaint against Zubair with the National Council for Protection of Child Rights (NCPCR), which wrote to Delhi and Raipur police alleging that Zubair's tweet amounted to "harassing and torturing" a minor girl online. Based on this, two FIRs were registered against Zubair, one in Delhi and one in Chhattisgarh under Sections 67 and 67A of the IT Act. The provisions of the Protection of Children from Sexual Offences Act and Section 509B IPC (sexual harassment online) were also reportedly invoked in Chhattisgarh.[31] Zubair filed a writ petition in the Delhi High Court for quashing the FIR and seeking protection against coercive action. The NCPCR acted on Jagdish Singh's complaint and summoned Twitter India while seeking action against Zubair.

The Delhi High Court granted interim protection to Zubair and directed Twitter to cooperate with the Delhi police and expedite the replies to their queries.[32] A month later, the Chhattisgarh High Court also granted interim protection to Zubair.[33] In both cases, the criminal proceedings were not quashed, even though it is difficult to imagine how his tweet was obscene, sexually explicit, or constituted sexual harassment or the "harassing and torturing" of the minor girl online, especially given that her photo had been blurred by him.

Zubair's case is an illustration of the misuse of Section 67 and 67A of the IT Act to target and criminalise political speech and activism in the name of protection of women. However, this case is not a one-off. In 2019, journalist Prashant Kanojia was arrested by the Uttar Pradesh police for offences under Section 67, IT Act and Section 500 and 505 IPC for allegedly making "objectionable

[30] For instance, in April 2020, when Zubair shared an old image of him wearing traditional Indian attire on Twitter, one Jagdish Singh replied to the post, commenting "once a jihadi is always a jihadi". In July, in response to another tweet posted by Zubair, Jagdish Singh replied "*Tu toh bada madharc**d nikla re*" (You have turned out to be a mother f***er) *(Ed. Note: obscured for publication)*. In August, when Zubair retweet an ALT news fact-checking report, Jagdish Singh once again abused him on Twitter.

[31] Karan Tripathi, "Delhi HC grants interim protection from arrest to Alt News co-founder Mohammed Zubair in plea seeking quashing of FIR", Livelaw, 9 September 2020, https://tinyurl.com/yxcafaqv; Livelaw, "Chhattisgarh HC grants interim protection from arrest to Alt News co-founder Mohammed Zubair in plea to quash FIR", 05 October 2020, https://tinyurl.com/y2oj5pef

[32] *Mohammed Zubair v State of GNCTD & Ors*, WP (Crl) No. 1429/20, order dated 09.09.2020, available at https://www.livelaw.in/pdf_upload/pdf_upload-381232.pdf

[33] "Altnews cofounder granted interim protection from arrest in harassment", Scroll, 05 October 2020, https://tinyurl.com/y5eqocj4

comments" against the U.P. Chief Minister Yogi Adityanath. On his Twitter account, Kanojia had shared a video news clip of a woman giving an interview to journalists claiming that she had sent a marriage proposal to Yogi Adityanath after talking to him for over a year. Kanojia shared this video with the comment "*Ishq chupta nahi chupaane se Yogi ji*" (roughly translated as "love cannot be hidden").[34] The police immediately filed an FIR under Section 500, IPC and Section 66, IT Act. Later, presumably realising that the sections were completely untenable – given that the offence of defamation under Section 500 IPC is a non-cognisable, bailable offence – the police issued a 'press note', adding Section 67, IT Act and Section 505, IPC, on the ground that Kanojia made "obscene comments" and "spread rumours" on social media, and arrested him. Although the Supreme Court eventually directed the release of Kanojia, they refused to quash the FIR,[35] even though there is nothing remotely objectionable about a video of a woman speaking to journalists.

> Not only are Sections 67 and 67A misused, the police also register cases of violation of consent and privacy of women as offences of obscenity under Section 67, IT Act, further erasing the harm experienced by women in such cases.

Even Section 66A continues to be used by the police, although it has been declared unconstitutional by the Supreme Court, including in Priyanka Sharma's case.[36] Sharma is a BJP worker who had shared an image of the West Bengal Chief Minister Mamata Banerjee's face superimposed on the actor Priyanka Chopra's MET Gala photograph. She was arrested by the police for the commission of offences under Sections 66A, 67A, and 500 IPC, based on a complaint by a Trinamool Congress party worker. She was released on bail by the Supreme Court only after she had tendered a written apology. It is unclear why the police invoked any of these provisions, especially given that there was nothing 'sexually explicit' about the image shared.[37]

In recent years, Sections 67 and 67A have also been used to register FIRs against a woman for sharing an "objectionable post" about UP CM Yogi Adityanath; Ekta Kapoor for alleged objectionable content

[34] Prashant Kanojia, https://twitter.com/PJkanojia/status/1136590104279937029

[35] Ankur Taliyan, "Decoding Section 67 of the IT Act, Section 505 IPC used in UP to arrest Prashant Kanojia", TimesNow, 11 June 2019, https://tinyurl.com/y5rbhnsp

[36] See Abhinav Sekhri and Apar Gupta, Section 66A and Other Legal Zombies, IFF Working Paper No. 02/2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3275893

[37] Vrinda Bhandari, "Mamata meme case: SC order discourages satire and free speech", Hindustan Times, 16 May 2019, https://www.hindustantimes.com/editorials/mamata-meme-case-sc-order-discourages-satire-and-free-speech/story-0XM5Nak9yZNabZdkwpUWDN.html; Deeksha Bharadwaj, "Why BJP worker Priyanka Sharma's arrest over Mamata meme is unlikely to hold up in court", The Print, 13 May 2020, https://theprint.in/india/governance/why-bjp-worker-priyanka-sharmas-arrest-over-mamata-meme-is-unlikely-to-hold-up-in-court/234855/

in the web series XXX Season 2 playing on her OTT platform AltBalaji; and Poonam Pandey and her husband in respect of an allegedly obscene video clip shot in Goa.[38]

Not only are Sections 67 and 67A misused, the police also register cases of violation of consent and privacy of women as offences of obscenity under Section 67, IT Act,[39] further erasing the harm experienced by women in such cases. Under the broadly worded provisions of Section 67, even the consensual sharing of a nude selfie by a woman with her partner is an offence, although statistics on whether, and how often, Section 67 is being used to register such cases are not available. If the partner non-consensually shares this image online (colloquially referred to as 'revenge porn'), it is a violation of privacy and bodily integrity caused by the non-consensual sharing of her image with the whole world.[40] However, in many instances, the police has treated such non-consensual sharing of sexual images online as an offence under Section 67/67A, instead of focusing on the offence of violating the privacy of the woman under Section 66E, IT Act.[41]

All this should give us some cause for concern before we advocate for *new* criminal laws.[42] Before the introduction of any new offence, we need to ask ourselves three questions. First, are the current laws inadequate to address the act that is sought to be criminalised, or do they in any manner fail to take into account the harm experienced by women? Second, what are the wordings of the proposed new offence? In many cases, the proposed offence may be worded in a vague and over-broad manner that, while intended to stop an illegal act, promotes misuse. Third, are there other legal and non-legal alternatives that may better address the online hate speech targeted at women?

[38] PTI, "Bengaluru woman booked for posting 'Objectionable Content' about Yogi Adityanath on Facebook", HuffPost, 21 March 2017, https://www.huffpost.com/archive/in/entry/bengaluru-woman-booked-for-posting-objectionable-content-about_a_21904458; PTI, "HC declines to quash FIR against Ekta Kapoor over web series", IndiaTV News, November 12, 2020, https://www.indiatvnews.com/entertainment/celebrities/hc-declines-to-quash-fir-against-ekta-kapoor-over-web-series-664362; TNN, "Poonam Pandey, hubby arrested in obscene video case, both get bail", Times of India, 06 November 2020, http://timesofindia.indiatimes.com/articleshow/79071041.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst

[39] Dutta, supra note 29.

[40] Vrinda Bhandari and Anja Kovacs, "What's sex got to do with it: Mapping the impact of questions of gender and sexuality on the evolution of the digital rights landscape in India" (CYRILLA Report, 2020).

[41] Dutta, supra note 29.

[42] We recognise that mere apprehension or possibility of misuse is not a reason for the State to refrain from criminalising a certain act (e.g., merely because the Protection of Women from Domestic Violence Act may be misused by women in some instances, does not mean that the law should not exist). However, care must be taken before advocating for new criminal laws in view of the value of the countervailing interests and the fact that alternative less intrusive measures could possibly be adopted. Further, any steps towards criminalisation must also consider what safeguards can be implemented to limit misuse of the law. We discuss some of the alternative measures and possible problems and safeguards in subsequent sections of this paper.

Other Legal Alternatives

Jurisdictions around the world are struggling with similar problems as in India. Many are indeed adopting methods of criminalisation and imposing greater obligations on intermediaries in this regard.[43] However, as discussed above, such methods (and indeed even applying narrowly tailored laws) may not always lead to optimal regulatory outcomes. What may be required, therefore, appears to be a more 'co-regulatory' system, which may temper the problems associated with excessive state intervention and platform inaction in a risk-based, proportionate manner.[44] In this context, two interventions bear further discussion.

The UK's Online Harms White Paper of 2019 is a holistic attempt at making the digital ecosystem safer, including for women and children.[45] Noting that extant systems of regulation fail to provide clear standards of what constitutes harm in the online ecosystem, the lack of clarity on how platforms should address harms, and the absence of a clear redress system for users, the report suggests making appropriate amendments to the law, based on evidence of harms. Importantly, it suggests casting a duty of care on platforms to take more responsibility for user safety, and provides measures to improve enforcement and transparency/accountability by platforms. Interestingly, it seeks to implement a 'safety by design' approach, similar to the concept of 'privacy by design', implying that user safety must be an intrinsic part of digital systems and not merely a feature.[46]

Germany's NetzDG law takes a slightly different approach by focusing largely on improving takedown processes and mandating increased transparency by large platforms.[47] While the law may not be as effective as desired, studies point to certain specific areas of improvement. These include the need to

[43] For instance, Australia amended its criminal code in 2019 to require a range of intermediaries to expeditiously remove "abhorrent violent material'' and notify law enforcement authorities within specific time periods.

[44] See for example Tiffany Li, "Beyond Intermediary Liability: The Future of Information Platforms", Workshop Report, Yale Law School, February 13, 2018, https://tinyurl.com/y4tm8dcg. Also see Bahl, Bailey and Rahman, supra note 4 and Article 19, "Self-regulation and hate speech on social media platforms", 2018, https://tinyurl.com/yxbnf5lt

[45] Note that the Report itself is not without criticism, including on the grounds that the recommendations may significantly limit speech rights. The Government of the UK is engaged in further consultations around the issue. Refer to "Online Harms White Paper", Secretary of State for Digital, Culture, Media, and Sport and the Home Department, Government of UK, April 2019, https://tinyurl.com/y4lnxxo9, "Mozilla submission to the UK government online harms whitepaper consultation", Mozilla EU Policy, last accessed on October 11, 2020, https://tinyurl.com/y29jqhal and Eline Chivot, "Online harms whitepaper – Some consultation responses", Inform Blog, July 20, 2019, https://inforrm.org/2019/07/20/online-harms-white-paper-some-consultation-responses/

[46] The White Paper recommends adoption of a risk-based and proportionate framework of enforcement, reliant primarily on industry created codes of conduct, which can be enforced by an independent regulator.

[47] Under the law which came into effect in 2018, platforms can be fined up to EUR 50 million for systemic failures to remove access to illegal content. It provides that companies must put in place easy-to-use complaints mechanisms, act to remove information within short time periods, and justify action taken in response to complaints. The law also recognises the use of "institutions of regulated self-governance" as dispute resolution forums. It also mandates reporting requirements. "Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act, Netz DG) – Basic Information, 2017", German Ministry of Justice, last accessed on October 11, 2020, https://tinyurl.com/y46t8c67

standardise complaints practices across platforms, create a 'clearing house' for complaints,[48] and to ensure greater independence and effectiveness of self-regulatory practices followed by platforms.[49]

In India too, it may be useful to focus on implementing co-regulatory frameworks aimed at improving procedures and enhancing transparency/accountability of platforms, rather than merely seeking to use blunt force instruments that criminalise certain types of content or seek to bar access to specific platforms.[50] It is worth emphasising again that process improvements must be mirrored at the level of the criminal justice system as well.

Interventions by the law could also seek to adopt a broader approach, say by creating more competition in the online ecosystem, including by enabling easier switching of services. Network effects and the presence of online monopolies can often put users in a dilemma – if they chose to forego the use of a particular platform (say due to the inadequate enforcement of hate speech policies), they may altogether lose the means to communicate with a large global audience.[51] The absence of adequate competition in the digital economy prevents the creation of feedback loops, which are an essential driver of consumer-oriented change in a market economy. Enabling easier switching of services will allow individuals to 'speak with their feet' – i.e., they can choose to use platforms with content moderation policies and practices that better serve their interests. In order for a platform to retain their user-base, they would therefore have to ensure their practices stay up-to-date with developing social norms.

[48] Which is said to enable easier pathways to dispute deleted content, provide barriers against over-blocking, and help keep companies accountable. William Echikson and Olivia Knodt, "Germany's NetzDG: A key test for combating online hate", CEPS Research Report No. 2018/09, November 2018, https://tinyurl.com/y57826dv

[49] William Echikson and Olivia Knodt, "Germany's NetzDG: A key test for combatting online hate", CEPS Research Report No. 2018/09, November 2018, https://tinyurl.com/y57826dv; and Article 19, "Self-regulation and hate speech on social media platforms", 2018, https://tinyurl.com/yxbnf5lt

[50] In fact, such procedural frameworks have indeed been adopted in India, notably in the context of protecting users of e-commerce platforms. Refer to the Consumer Protection (E-Commerce) Rules, 2020, https://consumeraffairs.nic.in/sites/default/files/E%20commerce%20rules.pdf

[51] See Omkar Khandekar, "Is Mastodon the messiah we need?", Mint, November 28, 2019, https://tinyurl.com/vb3ag4b, Deepak Gopalakrishnan, "Mastodon is temporary, Twitter is permanent: Why social media networks aren't easily replaced", Arre, November 9, 2019, https://tinyurl.com/y5b3ouq5, Aditi Agrawal, "Is Mastodon the answer to online hate speech?", MediaNama, November 8, 2019, https://www.medianama.com/2019/11/223-mastodon-online-hate-speech/. Also refer Diana Zulli, Miao Liu and Robert Gehl, "Rethinking the "social" in "social media": Insights into topology, abstraction, and scale on the Mastodon social network", New Media and Society, 22(7), pp.1188-1205, July 22, 2020, https://journals.sagepub.com/doi/10.1177/1461444820912533; and Adi Robertson, "How the biggest decentralised social network is dealing with its Nazi problem", The Verge, July 12, 2019, https://tinyurl.com/y3o83lka

Alternatives to 'Law' Based Methods

There is also a need to focus on alternative methods of regulation and adopt a more holistic approach to regulatory interventions in the digital ecosystem.[52] In the Indian context, it may also be preferable to avoid solutions that require excessive state intervention, not least due to poor rule of law frameworks, among other structural/systemic problems described above. One should therefore seek to examine the broader modalities of regulation that are available – including the use of social norms, the market, and technology, each of which, used correctly, can be leveraged to enable a safer online ecosystem.

For instance, pressure brought to bear on platforms by advertisers,[53] civil society or indeed employees of technology companies,[54] has proven (at times) to be effective in forcing positive change in a platform's content moderation practices. Multi-stakeholder initiatives, such as the development of the Santa Clara Principles in 2018, have also helped create global minimum standards for platforms to adhere to.[55] Such initiatives can help inform the development of legal standards, while also filling gaps in legislative efforts.

Using autonomy-preserving techno-regulatory methods such as nudges should be encouraged, with the necessary caveats pertaining to transparency and accountability.[56] For instance, the Scunthorpe

[52] This is particularly true due to the fast-changing nature of the digital ecosystem, the cross-jurisdictional aspects involved in online regulation, and the role played by the private sector in developing and governing this space.

[53] Most recently, over 800 advertisers from around the world withdrew from Facebook to protest against its hate speech policies. See Sanacy Scola, "Inside the ad boycott that has Facebook on the defensive", Politico, July 3, 2020, https://tinyurl.com/y8khl5ru, and Elizabeth Dwoskin and Taylor Telford, "Facebook is working to persuade advertisers to abandon their boycott. So far, they aren't impressed", Washington Post, July 4, 2020, https://tinyurl.com/y77y5vve. Note however, that Mark Zuckerberg, has at least in public, refused to bow down to pressure from advertisers to reform Facebook's policies. Tyler Sonnemaker, "Mark Zuckerberg reportedly said Facebook is 'not gonna change' in response to a boycott by more than 500 advertisers over the company's hate-speech policies", Business Insider, July 2, 2020, https://tinyurl.com/y2ocs2jh

[54] Therese Poletti, "Many tech employees are more 'woke,' but changing monolith companies is not easy", Market Watch, June 9, 2020, https://tinyurl.com/yyrqe4o8; Jeff Horowitz and Newley Purnell, "Facebook staff demands policy changes on India hate speech", The Wall Street Journal, August 21, 2020, https://tinyurl.com/y4jr34c3; Sobia Khan and Anumeha Chaturvedi, "Post hate speech controversy, Facebook conducts town hall with Indian employees", Economic Times, August 26, 2020, https://tinyurl.com/yytojcmt, Paige Leskin, "We would walk out with you — if Facebook would allow it: Content moderators join Facebook employees in revolt over how the company handles Trump posts", Business Insider, June 8, 2020, https://tinyurl.com/yypnh29r; Emma Graham-Harrison, "Rebels within: the Facebook staff openly challenging Zuckerberg", The Guardian, October 4, 2020, https://www.theguardian.com/technology/2020/oct/04/rebels-within-the-facebook-staff-openly-challenging-zuckerberg; Nitasha Tiku, "Three years of misery inside Google, the happiest company in tech", Wired, August 13, 2019, https://www.wired.com/story/inside-google-three-years-misery-happiest-company-tech/

[55] The Santa Clara Principles on Transparency and Accountability in Content Moderation, last accessed on October 11, 2020, https://santaclaraprinciples.org/. Also refer to Press Release, "One year after the release of the Santa Clara principles, OTI continues to push tech companies for transparency and accountability around content moderation practices", Open Technology Institute, May 7, 2019, https://tinyurl.com/y2yhw728; and Spandana Singh, "Assessing YouTube, Facebook and Twitter's content takedown policies", Open Technology Institute, May 7, 2020, https://tinyurl.com/yxzoongw

[56] See Simon Thompson, "Hate speech and self-restraint", Ethical Theory and Moral Practice 22, 2019, https://link.springer.com/article/10.1007/s10677-019-10004-y; and Karen Hao, "These simple design tricks can help diminish hate speech online", Quartz, October 11, 2017, https://qz.com/1093804/these-simple-design-tricks-can-help-diminish-hate-speech-online/

Sans font recognises and censors commonly-used swear words.[57] If made a default setting, using such a font could prove helpful in both preserving user autonomy while reducing hate speech.[58] Similarly, systems can be designed to recognise and warn users against the most egregious instances of hate speech, working in a manner similar to spell-checks. These need not completely bar the use of any particular text, but can serve to alert users to possibly problematic speech.[59]

Ensuring that these alternative modalities of regulation are designed to be non-coercive and developed through transparent and inclusive processes is important. Many of the non-legal modalities of regulation we see being implemented today are focused on or arise in the Global North. Involvement of stakeholders from developing countries must be strengthened.[60] The perspectives of multiple stakeholders, including from women and minorities, must be considered when designing or managing systems. This can be important not just in terms of creating a more just and efficacious regulatory framework, but also to prevent democracy deficit, which is inherent in non-state-led regulatory processes.

Conclusion

Online hate speech is a problem, especially since it is disproportionately targeted at women. While a law criminalising verbal abuse or sexual harassment online may seem attractive, past experience (with Section 66A and 67, IT Act and Section 509B, IPC) demonstrates that it will likely have unintended consequences and be misused, for instance in the form of restricting online expressions of autonomy and sexuality by women. We need to adopt a holistic approach that considers all interests, and reflects a balance of legal and non-legal alternatives that are not solely focused on criminalisation.

[57] "Scunthorpe Sans", vole.wtf, last accessed on October 11, 2020, https://vole.wtf/scunthorpe-sans/

[58] Various other technology-based initiatives include the use of 'blackbots', which allow users to share lists of blocked people amongst communities, and tools such as heartmob and trollbusters, that aim to engage the online community to aid victims of abuse. Susan Benesch, "Civil society puts a hand on the wheel: Diverse responses to harmful speech", Perspectives on Harmful Speech Online, Berkman Klein Center for Internet and Society, Harvard, August 2017, https://cyber.harvard.edu/sites/cyber.harvard.edu/files/2017-08_harmfulspeech.pdf

[59] In the context of privacy policies, research indicates that giving people time to 'cool off'' leads to better understanding of the terms they read. Similar principles could be applied to either explain to users the nature of permitted conduct, or to provide warnings and opportunities to retract comments before publication, thereby allowing users to reconsider abusive behaviour. IntAct. "Cool down for better understanding", accessed on October 12, 2020, https://intactprivacy.com/cool-down-for-better-understanding/

[60] For instance, Twitter recently banned President Trump from their platform and Google and the Apple App Store took down Parler, for concerns of hate speech and incitement to violence. Regardless of one's views on the desirability of such an action by a private platform, it is important to note that cases of hate speech are never regulated with such vigour in developing countries like India. Similarly, Instagram changed its nudity policy in 2020, following outcry over its decision to censor semi-nude empowerment pictures posted by women who are plus-size and women of colour. While recognising the merits of such a campaign, it is questionable whether a similar outcry in India would have borne fruit.

> As UNESCO notes, focusing primarily on repressive criminal measures may misunderstand the complexity of the situation, and miss out on an opportunity for coordinated and holistic responses from all the stakeholders in society.

As UNESCO notes, focusing primarily on repressive criminal measures may misunderstand the complexity of the situation, and miss out on an opportunity for coordinated and holistic responses from all the stakeholders in society.[61] The Criminal Law Reform Committee constituted by the Indian government in May 2020 – tasked with overhauling the substantive and procedural criminal laws in the country – provides us with a timely opportunity to make the law work for women and all other vulnerable groups. However, as long as the Committee continues to comprise only of Hindu, upper-caste men, an honest conversation is not possible.

---

[61] UNESCO, Countering Online Hate Speech, 2017, https://tinyurl.com/yypnkm85