



RESPONSE TO CALL FOR SUBMISSIONS ON ‘CONTENT REGULATION IN THE DIGITAL AGE’¹

Centre for Communication Governance, National Law University Delhi

We thank the United Nations Special Rapporteur on Freedom of Expression, David Kaye for inviting comments for the June 2018 Human Rights Council Report on ‘Content Regulation in the Digital Age’. This note makes reference to some of our other work, which we will be sharing along with this paper.

We begin with a brief background note about freedom of expression and online intermediaries in India. After this we deal with two of the questions listed in the Call for Submissions.

Background note

The right to freedom of expression in India is codified in Article 19(1)(a) of the Indian Constitution. It has been read to include press freedom,² and to extend to speech on the internet.³

General laws criminalising specified categories of speech apply online in India. For example, Section 3⁴ of the Scheduled Caste and Scheduled Tribe (Prevention of Atrocities) Act, 1989 (‘SC/ST Act’): the Delhi High Court confirmed in *Gayatri vs. State*,⁵ that derogatory casteist

¹ With inputs from Chinmayi Arun, Arpita Biswas and Aditya Singh Chawla from the Centre for Communication Governance at National Law University Delhi

² Romesh Thapar vs. State of Madras, 1950 SCR 594; Sakal Papers vs. Union of India, AIR 1962 SC 305; Bennett Coleman vs. Union of India, AIR 1973 SC 106.

³ Shreya Singhal vs. Union of India, (2013) 12 SCC 73.

⁴ 3. Punishments for offences of atrocities--

(1) Whoever, not being a member of a Scheduled Caste or Scheduled Tribe,--

...

(x) intentionally insults or intimidates with intent to humiliate a member of a Scheduled Caste or a Scheduled Tribe, in any place within public view:

...

shall be punishable with imprisonment for a term which shall not be less than six months but which may extend to five years and with fine.

⁵ *Gayatri vs. State*, W.P.(CRL) 3083/2016, available at <https://barandbench.com/wp-content/uploads/2017/07/sc-st-judgment-watermark.pdf> (Last visited on February 1, 2018).

slurs posted on Facebook can be punishable under Section 3(1)(x) of the SC/ST Act. This implies that Section 3 of the SC/ST Act can be applicable to other social media platforms.⁶

The Indian government derives its power to block online content from Section 69A of the Information Technology Act, 2000 ('IT Act') read with the Information Technology (Procedure and Safeguards for Blocking for Access of Information by Public) Rules, 2009 ('Blocking Rules').⁷ Although the Indian Telegraph Act, 1885 also empowers the Indian government to block content, we are not discussing it in this note since the statute only applies to 'telegraphs' and requires the presence of physical apparatus.⁸ The Telegraph Act is relevant in the context of content blocking by internet service providers and other kinds of communication service providers.⁹

The Indian Supreme Court first discussed freedom of expression in the context of web-based intermediaries in *Shreya Singhal vs. Union of India*.¹⁰ It ruled that freedom of expression applies online, and read down the notice and take down rules¹¹ such that intermediaries were no longer required to judge which private notices for content removal were legitimate. This notice and take down system resulted from the IT Act: the immunity from liability (or 'safe harbour') offered to intermediaries¹² was subject to certain conditions,¹³ one among which was that intermediaries had to remove illegal content upon notice.¹⁴ After *Shreya Singhal v. Union of India*, intermediaries are only required to remove content if required to do so by a court order, or if notified by the appropriate government or its agency.¹⁵

However, the jurisprudence on the intermediaries' liability for the content they host and transmit remains precarious. For example, *Sabu Mathew George vs. Union of India*,¹⁶ was about search engines displaying content that violated the Pre-Conception and Pre-Natal Diagnostic Techniques (Prohibition of Sex Selection) Act, 1994 ('PNDT Act').¹⁷ The

⁶ Abhinav Garg, *Social Media slurs on SC/ST punishable: HC*, THE TIMES OF INDIA, (July 4, 2017) available at <http://timesofindia.indiatimes.com/india/social-media-slurs-on-sc/st-punishable-hc/articleshow/59432794.cms> (Last visited on February 1, 2018).

⁷ Our report on intermediary liability titled 'NOC Online Intermediaries Case Studies Series: Online Intermediaries in India' discusses the content blocking mechanism and its opacity in detail.

⁸ Section 3(1AA), Indian Telegraph Act, 1885.

⁹ We discuss the regulatory framework in detail in our India report for the Freedom on the Net report, 2017.

¹⁰ *Shreya Singhal vs. Union of India*, (2013) 12 SCC 73.

¹¹ The Information Technology (Intermediary Guidelines) Rules, 2011.

¹² Section 79, Information Technology Act, 2000.

¹³ Rule 3, The Information Technology (Intermediary Guidelines) Rules, 2011.

¹⁴ Our report titled 'NOC Online Intermediaries Case Studies Series: Online Intermediaries in India' discusses the Intermediary Guidelines in further detail.

¹⁵ *Shreya Singhal vs. Union of India*, (2013) 12 SCC 73, at para 119.

¹⁶ *Sabu Mathew George vs. Union of India*, W.P.(C) No. 341/2008.

¹⁷ Section 22, Pre-Conception and Pre-Natal Diagnostic Techniques (Prohibition of Sex Selection) Act, 1994. "Prohibition of advertisement relating to pre-natal determination of sex and punishment for contravention. (1) No person, organization, Genetic Counselling Centre, Genetic Laboratory or Genetic Clinic shall issue or cause to be issued any advertisement in any manner regarding facilities of pre-natal determination of sex available at such Centre, Laboratory, Clinic or any other place. (2) No person or organization shall publish or distribute or cause to be published or distributed any advertisement in any manner regarding facilities of pre-natal determination of sex available at any Genetic Counselling Centre, Genetic Laboratory, Genetic Clinic or any other place."

petitioners sought a ruling requiring pre-censorship or ‘auto-blocking’ of the content under question by the online intermediaries, using key words if necessary. The search engines argued that they should only be required to remove content specifically flagged by the government nodal agency (created for this purpose). Eventually a government nodal agency was set up, to review and forward complaints about advertisements for sex identification to the search engines. The search engines in turn were required to set up expert bodies that would delete anything that violated the PNDA Act.

Similarly, in *In Re: Prajwala Letter*¹⁸, which is still being heard by the Indian Supreme Court, discusses videos of sexual violence being distributed on the internet. As a part of this case, a committee was constituted to advise the Court on the “feasibility of ensuring that videos depicting rape, gang rape and child pornography are not available for circulation.” The committee consisted of government experts, industry representatives and lawyers involved with the case,¹⁹ and it heard a number of technical experts before preparing its report but did not seek input from any organisation or individual working on freedom of expression in India.²⁰ The report has not been made public.

There appears to have been consensus within the advisory committee on the creation of a government controlled hash bank of rape and gang rape content.²¹ There also appears to have been consensus that independent sites and applications that host child pornography and rape or gang rape content, and do not ‘remove such contents of their own’ can be blocked by law enforcement agencies.²² Although the judgment in this case is still pending, the hearings have resulted in a series of orders that suggest that the global companies are working with the Indian government to build an opaque content blocking mechanism. Although the content that is the subject of this particular case is clearly harmful and must be blocked, the content blocking procedure does not seem to have built in any safeguards to prevent misuse, or offer redress when over-blocking takes place.

CCG Response to Specific Questions in the Call for Submissions

Q. 3 – Individuals at risk: *Do company standards adequately reflect the interests of users who face particular risks on the basis of religious, racial, ethnic, national, gender, sexual orientation or other forms discrimination?*

In our opinion, the company standards do not reflect the interests of individuals at risk, especially in contexts in which risk is local or hyper-local.²³ In part (a) we explain why the manner in which company standards have been framed is inherently problematic from the point of view of individuals at risk; and in part (b) we discuss context and hyper-local harm, with

¹⁸ *In Re: Prajwala Letter Dated 18.2.2015 Videos of Sexual Violence and Recommendations*, SMW (CrI) No. 3 of 2015 (Supreme Court of India).

¹⁹ Order dated October 23, 2017, *In Re: Prajwala Letter Dated 18.2.2015 Videos of Sexual Violence and Recommendations*, SMW (CrI) No. 3 of 2015, (Madan B. Lokur and Uday Udesch Lalit, JJ.) (Supreme Court of India).

²⁰ *Id.*

²¹ *Id.* at 8.

²² *Id.* at 12.

²³ Chinmayi Arun, *Rebalancing regulation of speech: Hyper-local content on global web-based platforms*, (January 23rd, 2018) available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3108238.

illustrations of how the standards can fail to protect particular kinds of vulnerable groups.

- (a) The major global web-based platforms have what might be described as broadly framed standards in place. For example, Twitter's policies forbid incitement of harm, promotion of violence, threats and direct attacks on people based on race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or disease.²⁴ Facebook prohibits hate speech in the form of content that attacks people based on their actual or perceived race, ethnicity, national origin, religion, sex, gender or gender identity, sexual orientation, disability or disease.²⁵

However, information in the public domain suggests that no distinction is made between vulnerable and powerful groups within these categories. As the *ProPublica* investigation²⁶ has demonstrated, this can lead to 'white men' being treated the same as 'black men', and to a religious majority being treated the same as a persecuted religious minority. This assumes that speech against vulnerable groups and powerful groups is equally harmful, and creates risks for members of vulnerable groups who speak out against discrimination.

The opacity of the review mechanism and the lack of accountability for incorrect application of the standards might even lead to disproportionate censoring of opinions of minority communities, or permit violent speech directed against them.

- (b) The standards do not account for all vulnerable groups. For example, we are given to understand that the companies do not interpret 'race' to cover caste, although violence directed at vulnerable caste groups has been a serious concern in India all through its history. The Indian government has resisted the idea that race must be interpreted to include caste,²⁷ although activists have argued otherwise.²⁸ The 2016 report of the UN Special Rapporteur on minority issues has however recognised discrimination based on caste and 'analogous systems of inherited status'.²⁹ Given the global nature of the web-based platforms, it may be advisable for them to consider a similarly expansive definition of protected categories of people in their definition of hate speech.

²⁴ Twitter, *Hateful Conduct Policy*, available at <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (Last visited February, 2018).

²⁵ Facebook, *Community Standards, Encouraging Respectful Behaviour - Hate Speech*, available at <https://www.facebook.com/communitystandards#hate-speech> (Last visited on February 1, 2018).

²⁶ Julia Angwin and Hanges Grassegger, *Facebook's Secret Censorship Rules Protect White Men from Hate Speech but not Black Children*, PROPUBLICA (June 28, 2017) available at <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms> (Last visited on February 1, 2018).

²⁷ Naunidhi Kaur, *Caste and Race*, 18(3) FRONTLINE, (2001) available at <http://www.frontline.in/static/html/fl1813/18130950.htm> (Last visited on February 1, 2018).

²⁸ *Id.*; Abdullah Nasir and Priya Anuragini, *India Must Not Shield Itself from International Scrutiny on Caste Discrimination*, THE WIRE, (December 26, 2017) available at <https://thewire.in/203883/india-must-not-shield-itself-from-international-scrutiny-on-caste-discrimination/> (Last visited on February 1, 2018).

²⁹ United Nations General Assembly, *Report of the Special Rapporteur on Minority Issues*, (January 28, 2016) available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G16/013/73/PDF/G1601373.pdf?OpenElement> (Last visited on February 1, 2018).

We want to note that there is likely to be some difficulty in understanding and implementing the standards, even if they are expanded. Some terms can have hyper-local meaning and can result in hyper-local harm.³⁰ This would be difficult to identify even by people living in the same country. This was evident from the ‘Piano Man’ incident in New Delhi, where the Indian owner of a jazz club did not realise at first glance that the title of an event being hosted at his club demeaned a vulnerable caste.³¹ This hyper-local harm on web-based platforms is likely to affect the most vulnerable groups, about whom there is little awareness and for whom there are not very many spokespeople.

The platforms’ standards on extremist content also fail to account for sensitivities and the delicate balance of human rights within a conflict zone: broad application of terms and conditions forbidding extremist content may be in the interests of national security but may also have a great impact on fragile political rights. In Kashmir for example, in the aftermath of the murder of Hizbul Mujahideen commander Burhan Wani in Kashmir, Facebook posts discussing the incident were being deleted. This included content shared by people residing in Kashmir as well as scholars in the United States and the United Kingdom. The content was removed for being in violation of ‘community standards’.³² However some of the speech removed was arguably politically valuable speech that ought to have been protected. The importance of permitting this speech on the platform becomes clearer if one considers that there tends to be heavy-handed press censorship in Kashmir, and this was very much in evidence after the Burhan Wani killing.³³

Q. 5 – Bias and Non-Discrimination: *How do companies take into account cultural particularities, social norms, artistic value, and other relevant interests when evaluating compliance with terms of service? Is there variation across jurisdictions? What safeguards have companies adopted to prevent or redress the takedown of permissible content?*

Cultural particularities and social norms are sometimes codified. We have discussed examples of this above – the SC/ ST Act criminalises certain kinds of speech targeting members of vulnerable caste groups, and is a response to a form of violence particular to the region. It appears that companies are not yet able to cope with these speech norms, even though they are codified.

The other example of a codified response to a regional cultural peculiarity, is the Pre-Natal Diagnostic Techniques (Regulation and Prevention of Misuse Act), 1994. This was enacted in

³⁰ Arun, *supra* note 23.

³¹ Shantanu David, *Delhi’s Piano Man Jazz Club Faces Music over ‘Offensive’ Band*, NEWS18 (June 9, 2017) available at <http://www.news18.com/news/lifestyle/food-delhi-restaurant-and-performance-space-faces-online-outrage-over-offensive-band-name-1424447.html> (Last visited on February 1, 2018).

³² Arun, *supra* note 24.

³³ Scroll Staff, *Kashmir press gagged’: Local newspapers react to the first formal press ban in the Valley*, SCROLL.IN (July 17, 2016) available at <https://scroll.in/article/811934/kashmir-press-gagged-local-newspapers-react-to-the-first-formal-press-ban-in-the-valley> (Last visited on February 1, 2018).

response to the practice of female foeticide and infanticide.³⁴ As a result, ‘advertisements’ relating to pre-natal sex determination are prohibited in India.³⁵ The global search engines did not take this into account, leading to *Sabu Mathew George vs. Union of India*,³⁶ to which Microsoft, Yahoo and Google were parties.

Separately, we are concerned about how far the companies engage with uncodified social norms that affect vulnerable groups’ speech and engagement online. For example, women face a high risk of violence in India and their sexuality is controlled closely. This may mean that an otherwise ordinary photograph – a young woman in jeans³⁷ or holding a drink – may place the subject at risk of violence depending on which part of India she is from³⁸. Real name policies may exacerbate the risk by making it easier to find and monitor the vulnerable person.³⁹

Other contexts in which it may be important to take social norms into account are when there is a political movement by a vulnerable group. In 2017, as a part of the #MeToo movement, Raya Sarkar published a crowdsourced list of scholars and teachers in India accused of sexual harassment by anonymous contributors on Facebook.⁴⁰ The list attracted criticism, and spurred a heated debate on the social media platform and beyond it. There was a caste dimension to the debate - many of the people who promoted and defended the list were members of the ‘Ambedkarite’ movement. In the middle of the debate, Sarkar’s account was blocked by Facebook for violating its terms and conditions. Her account was eventually restored through backchannel intervention. However, it illustrated the problem with Facebook’s system of appeals and remedies.⁴¹

The political significance of Raya’s speech at that particular time illustrates how over-broad

³⁴ Anuradha Mascarenhas, *Sex Determination: An Old Law, A New Debate*, THE INDIAN EXPRESS, (February 4, 2016) available at <http://indianexpress.com/article/explained/sex-determination-an-old-law-a-new-debate/> (Last visited on February 1, 2018); Ministry of Health and Family Welfare, *Annual Report on the Implementation of the Pre-Conception and Pre-Natal Diagnostic Techniques (Prohibition of Sex Selection) Act, 2006*, available at <http://pndt.gov.in/writereaddata/mainlinkfile/File99.pdf> (Last visited on February 1, 2018).

³⁵ Section 22, Pre-Natal Diagnostic Techniques (Regulation and Prevention of Misuse Act), 1994.

³⁶ *Sabu Mathew George vs. Union of India*, W.P. (Civil). No. 341 of 2008.

³⁷ Suresh Foujdar, *No mobile phones and jeans, khap panchayat tells girls in Rajasthan*, HINDUSTAN TIMES, (July 2, 2017) available at <https://www.hindustantimes.com/jaipur/no-mobile-phones-jeans-a-panchayat-in-rajasthan-tell-girls/story-rVVetLLsUYxvIVBHbPloL.html> (Last visited on February 1, 2018).

³⁸ Lalita Panicker, *Honour killings are products of social prejudice against women*, HINDUSTAN TIMES, (October 8, 2016) available at <https://www.hindustantimes.com/columns/honour-killings-are-products-of-social-prejudice-against-women/story-qPSWbpmoQ2NrOeF2qkkjPM.html> (Last visited on February 1, 2018).

³⁹ Melissa Chan, *Facebook Relaxes Controversial ‘Real-Name’ Policy After Criticism*, TIME, (December 15, 2015) available at <http://time.com/4150410/facebook-real-name-policy/> (Last visited on February 1, 2018); Sam Levin, *As Facebook blocks the names of trans users and drag queens, this burlesque performer is fighting back*, THE GUARDIAN, (June 29, 2017) available at <https://www.theguardian.com/world/2017/jun/29/facebook-real-name-trans-drag-queen-dottie-lux> (Last visited on February 1, 2018).

⁴⁰ Shreya Row Chowdhury and Aroon Deep, *Universities Respond to Raya Sarkar’s list of alleged sexual predators: Mostly silence, some denials*, SCROLL.IN, (November 4, 2014) available at <https://scroll.in/article/856589/universities-respond-to-raya-sarkars-list-of-alleged-sexual-predators-mostly-silence-some-denials> (Last visited on February 1, 2018).

⁴¹ Aditya Singh Chawla, *Facebook and its dis(contents)*, (January 8, 2018) THE CCG BLOG, available at <https://ccgnludelhi.wordpress.com/2018/01/08/facebook-and-its-discontents/> (Last visited on February 1, 2018).

enforcement of terms and conditions without an effective and speedy appeals mechanism can affect politically important, controversial speech. However this is complex since the same naming and shaming exercise would have been unacceptable if directed by a powerful group at a vulnerable group of people. For example, a list put together by the powerful castes and targeting scheduled caste scholars would have very different implications. This is why a neutral standard (such as race or caste) that does not take power and vulnerability into account may prove counter productive.