EVENT REPORT

How Safe is Your Harbour? Discussions on Intermediary Liability and User Rights

July 10, 2020

By **Anna Liz Thomas**, **Shubhika Saluja**, **Kanav Khanna** and **Dipto Ghosh**
Edited by **Gurshabad Grover** and **Torsha Sarkar**

The Centre for Internet and Society, India

The event was supported by the John D. and Catherine T. MacArthur Foundation and Facebook India.

Introduction

On 10 January 2020, the Centre for Internet and Society (CIS) hosted panel discussions on intermediary liability in India. The event consisted of four panel discussions with the overarching focus being on intermediary liability in India, discussed primarily with reference to the Draft Intermediary Guidelines Rules that were released by the Ministry of Electronics and Information Technology for public consultation in December 2018 (the "Draft Rules").[1]

The first panel focussed on one of the more controversial and stringent rules introduced in the proposed amendments, Rule 3(9), which necessitates the use of automated technology by intermediaries for filtering 'unlawful' content. The draft rule does not specify the scope of the content to be detected, the technologies to be used, or any procedural safeguards that accompany the deployment of the technology. The key discussion points included the rationale behind the introduction of this mandate, the technical infeasibility of implementation, its effect on freedom of speech online, and possible mitigations.

The second panel was themed around content takedown and its legal aspects specifically under S.69 and S.79 of the IT Act, which permit the Government to mandate intermediaries to remove/block content. The panel discussed the ramifications of the flawed legal system that regulates content takedown, the failure of the intermediary liability regime to sufficiently accommodate issues relating to content removal, and ways of improving and introducing nuances in the legal framework.

The third panel looked at the proposal requiring intermediaries to enable traceability of originators of information. While this move is ostensibly to crack down on misinformation and fake news, there are questions regarding its feasibility and effects on platform architecture. More importantly, it poses dangers for the freedom of expression and privacy of users. These questions were juxtaposed against the ongoing litigation in the WhatsApp case. The discussion was also themed around the proposal made by IIT Madras which attempted to ensure traceability without breaking encryption, and the possible constitutional concerns emanating from the implementation of this rule.

The final panel brought together the threads from the previous discussions, and debated the ways in which the draft intermediary guidelines represent a departure from the current model of intermediary liability in India, and its potential effects on similar regulation in other countries.

[1] Draft Information Technology [Intermediaries Guidelines (Amendment)] Rules 2018 <https://meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf> accessed 17 June 2020

Automated Content Filtering

The first panel discussion revolved around the use of automated technology in order to proactively filter content, as proposed by Rule 3(9) of the Draft Rules.

Panelists

- Kanksshi Agarwal (Senior Researcher, Centre for Policy Research)
- Nayantara Ranganathan (Independent researcher)
- Shashank Mohan (Counsel, Software Freedom Law Centre)
- *Moderator:* Akriti Bopanna (Policy Officer, CIS)

Rule 3(9) in Context: Why was the rule introduced?

The panel opened with a discussion regarding the intention behind the inclusion of Rule 3(9) in the Draft Rules and the problems that it sought to address. Nayantara Ranganathan contextualised the Draft rules against the 'Calling attention motion in parliament on the need to address the misuse of social media platforms and the spread of fake news', suggesting that these rules were brought out in response to the problem cited. Kanksshi Agarwal highlighted national security and defamation as concerns for the current environment, leading to the government feeling the need for censorship of certain viewpoints, which further set the backdrop for the Rule. Shashank Mohan brought up the press release[2] by the Press Information Bureau when the Draft Rules were introduced. The release indicated the government's intention to act against the misuse of social media by 'criminal' and 'anti-national' elements, as well to prevent lynching incidents as a result of fake news circulating on WhatsApp.

Meaning of 'technology based automated tools'

The conversation then shifted towards what is meant by the phrase 'technology based automated tools' as used in Rule 3(9). Since the same has not been defined anywhere else in the Rules or the parent Act, the panel speculated various kinds of mechanisms that this phrase may refer to. Filtering could be done using more traditional simpler tools such as upload filters and metadata-based filtering or by using more advanced technologies such as artificial intelligence and machine-learning, which embed more context in the process.

The term 'automated tools', Shashank Mohan argued, should be contextualized within the various mechanisms that 'big tech' companies were already utilizing. If the government introduced Rule 3(9) to filter content like child pornography and terrorist content, categories of content that these companies were already filtering with a modicum of success, it was pertinent to consider the ends that the government was seeking to fulfil with this rule. Each technology had its own limitations as to the accuracy of what gets filtered out. Machine technology was not able to understand context and thus imposing a

[2] 'Draft IT rules issued for public consultation', *Press Information Bureau* (New Delhi, 24 December 2018) <https://pib.gov.in/PressReleasePage.aspx?PRID=1557159> accessed 17 June 2020

uniform metric for censorship often resulted in the unwarranted takedown of legitimate content.

Concerns emanating from Rule 3(9)

Fake News

The panel next deliberated on the nature of fake news, the addressal of which has been identified by the Government as a central reason for these Rules. Nayantara Ranganathan brought up how misinformation and disinformation is spread through the interaction of various kinds of content that form the information ecosystem, with State-backed agents being a major proponent of the same. In this context, fake news becomes a placeholder for all kinds of conversations about information disorder, but is also used strategically as a means to de-legitimize facts or statements that are contrary to a certain understanding of an issue. Cutting edge research on misinformation focuses not just on detecting what is false but also on the need for reinstating public trust and legitimacy to conventional social institutions such as the media and the courts. In this larger context, the act of proactive filtration seems to be a simplistic approach towards solving the issues that fake news has created. Shashank Mohan further debated whether fake news concerns were to be discussed in the context of the 'traceability' requirement of Rule 3(5) of the Draft Rules, or under the 'proactive identification' requirement under Rule 3(9).

Freedom of Speech

Concerns were raised by Shashank Mohan regarding the status of freedom of speech in India in the present political environment. The Supreme Court has, in the past, considered pre-censorship to be anathema to the right to free speech.[3] In that light, upload filters create entry barriers to the internet, and the marketplace of ideas. When the state dictates what should enter the marketplace of ideas, it would have rippling effects across civil liberties, democratic institutions and democracy at large. In a free speech regime, it should be presumed that all speech is legal and then certain restrictions should be placed as opposed to beginning with the assumption that all speech is forbidden and subsequently allowing content based on the status of the speaker. Automated filtering mechanisms would thereby impugn this freedom and lead to an increased marginalisation of already marginalised groups on the internet. Shashank Mohan brought up an instance of Facebook removing content for Indian viewers, but not for those located outside of India, to show how content removal was no longer simple. Self-regulation by companies in content takedown added further opacity to the process.[4] Another example that was cited by Kanksshi Agarwal, was the removal of evidence of

[3] See *R. Rajagopal vs. State of Tamil Nadu*, 1995 AIR 264
Also see Gautam Bhatia, 'Free Speech and Public Order' (*CIS India*, 17 February 2016) <https://cis-india.org/internet-governance/blog/free-speech-and-public-order-1> accessed 17 June 2020

[4] Ather Zia, 'How Facebook helps silence Kashmiri Voices', *The Caravan* (online) (1 December 2018) <https://caravanmagazine.in/commentary/how-facebook-helps-silence-kashmiris> accessed 17 June 2020

human rights violations in Syria by platforms under pressure to remove 'extremist content'.[5]

Accountability

Since content filtration can take place based on a company's own policies, as well as complying with whatever the government requires, there is a question of accountability that must also be addressed, since companies may not just be filtering content based on the restrictions set out in Article 19(2) of the Indian Constitution. Intermediaries are supposed to be neutral and not exercise editorial functions, but due to self-regulation measures and added pressure by states, a lot of intermediaries are making non-transparent algorithmic determinations on what should be seen by users. The panellists also discussed how editorial responsibility is indirectly being placed on intermediaries, and how, in that context, differential treatment of different intermediaries must be done for ensuring better accountability. By giving law enforcement and policing functions to the intermediary, the state was divesting some of its functions to the intermediary. In this context the intermediaries becomes the state's proxy to achieve state goals, which could become problematic especially given that traditionally speech rights are enforceable against the state, and not against private parties.

Practicality

The panellists discussed whether it was practically feasible for every kind of intermediary to incorporate automated tools given the resources, both financial and human, that would be required to put such mechanisms in place. Nayantara Ranganathan pointed out how a lot of what is pushed forward as AI based content moderation also requires a lot of manual training of the data and for tagging content. Despite this, even the companies with the most amount of resources using exploitative means are still not able to solve the problems. She also highlighted how the responses and public comments by some intermediaries to the Draft Rules, had indicated that users had a legitimate expectation that there would be humans involved in decision making regarding content, an expectation that is thwarted by the language in the Draft Rules.[6]

Transparency

To ensure transparency, Kanksshi Agarwal emphasized that the Santa Clara principles should be incorporated, which mandates a notice should be sent to users whose content is taken down, and they should be given the opportunity to appeal the said takedown if it is believed to be unwarranted. The panelists discussed the obligation of intermediaries to be more forthcoming to users when content gets taken down, the reasons for the same and the process of appeals for content-takedown, and providing a right to fair hearing.

---

[5] Hadi Al Khatib, Dia Kayyali, 'YouTube is Erasing History', *The New York Times* (October 2019) <https://www.nytimes.com/2019/10/23/opinion/syria-youtube-content-moderation.html> accessed 17 June 2020

[6] For example, see Change.org's Public Comment on the Draft Rules; Meity.gov.in, *Public Comments on Draft Intermediaries Guidelines Rules* (2018), p.57 <https://meity.gov.in/writereaddata/files/public_comments_draft_intermediary_guidelines_rules_2018.pdf> accessed 17 June 2020

Conclusion

The final subject of discussion was regarding the legitimacy of Rule 3(9), if appropriate safeguards were introduced, given the delegation of State duty to intermediary organisations. The practical reality of the current political environment would have to be considered to contextualise the rule, and the discomfort felt by people regarding the broad powers under the rule. Apart from governments and companies, the need for user voices to be included in decision making processes regarding online regulation of speech was highlighted.

Content Takedown

The second session was themed around content moderation and content regulation.

Panelists

- Bhavna Jha (Research Associate, IT for Change)
- Divij Joshi (Technology Policy Fellow, Mozilla)
- *Moderator*: Torsha Sarkar (Policy Officer, CIS)

Contextualising Content Moderation

The panel began with Divij Joshi setting three principled provocations of content moderation, which would serve as the broad contextual setting for the discussion. First, content moderation lies at the core of social media operation and not an ancillary aspect of their functioning.[7] Therefore to understand the issue of content moderation, we must understand the politics of social media. Second, content moderation frameworks are necessary. Several feminist scholars have pointed out that the concept of the internet as a 'marketplace of ideas' may be subversive to the rights and expression of marginalized communities.[8] Third, the intermediary liability framework was insufficient for regulating content moderation and regulation. While the former only determined the extent of an entity's criminal/civil liability, the latter raised more nuanced issues of setting standards for functioning. Intermediary liability and content moderation as ideas were linked, but insufficient to address the issues presented.

Bhavna Jha also laid down the contemporary context to better understand why a conversation on content moderation was necessary. Several reports[9] were mentioned which demonstrated the inability of social media platforms in addressing cyber abuse,

---

[7] See Tarleton Gillespie, 'Platforms are not Intermediaries' (2018) 2 GEO. L. TECH. REV. 198 <https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Gilespie-pp-198-216.pdf> accessed 17 June 2020

[8] See for example, the Gamergate controversy, where harassment campaigns were initiated against women in the video game industry on platforms like Twitter and Reddit where women were subjected to doxing, threats of rape and death threats; Emily VanDerWerff, 'Why everybody in the video game world is fighting' (*Vox*, 13 October 2014) <https://www.vox.com/2014/9/6/6111065/gamergate-explained-everybody-fighting> accessed 17 June 2020
In the Indian context, platform moderation policies have been accused of being casteist by taking arbitrary action against marginalized classes, while being soft on those who peddle hate speech. See Himani Chandna, 'Twitter caught in caste controversy again, users say it discriminates against SCs, STs and OBCs' (*The Print*, 4 November 2019)
<https://theprint.in/india/twitter-caught-in-caste-controversy-again-users-say-it-discriminates-against-sc-st-obcs/315546/> accessed 17 June 2020

[9] See Amnesty International, *Toxic Twitter: A Toxic Place for Violence Against Women* (2018), <https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/> accessed 17 June 2020;
IT for Change, *Born Digital, Born Free? A Socio-Legal Study on Young Women's Experiences of Cyberviolence in South India* (2019)
<https://itforchange.net/born-digital-born-free-a-socio-legal-study-on-young-womens-experiences-of-cyber violence-south-india> accessed 17 June 2020;
Equality Labs, Facebook India: *Towards the Tipping Point of Violence: Caste and Religious Hate Speech* (2019);
<https://static1.squarespace.com/static/58347d04bebafbb1e66df84c/t/5d0074f67458550001c56af1/1560311033798/Facebook_India_Report_Equality_Labs.pdf> accessed 17 June 2020

harassment and hate speech issues over the internet. Given that intermediaries currently function with relative impunity and arbitrariness when making decisions concerning content moderation and account deletion, this necessitates discourse on the responsibilities of intermediaries through publicly accountable content takedown mechanisms.

Problems with the Existing Regime

The discussion then moved onto the shortcomings within the existing regime on content moderation. In this context, panellists deliberated on the problems caused by the Shreya Singhal judgement. While it did strike down the broadly worded provision section 66A, it resulted in the absence of an effective mechanism for content takedown. In a country with low digital literacy and restricted access to justice, the current mechanism which requires an individual to approach a government agency or court for an order is impractical. It has empowered platforms to abdicate accountability leaving marginal communities vulnerable. This legal vacuum of accountability is increasingly being filled with arbitrary exercise of executive power.

Similarly, Divij Joshi argued that the regime under section 69A is insufficient in addressing content moderation. The rules under section 69A require a committee headed by a Designated Officer to examine and approve all requests relating to content takedown. It is natural that there would be very limited application of mind by the executive, when there hundreds of daily takedown requests to be approved. Further, Bhavna Jha pointed out how, content moderation which seems to be permitted under rule 3(8) and 3(9) of the Draft Rules, seems to be using the language of the reasonable restrictions on speech as set out in Article 19(2) of the Constitution. This is simply delegating to intermediaries and to police officers, the power to determine how to interpret Article 19(2)). Consequently, narrow and precise categories for takedown of 'unlawful' content must be introduced. Divij Joshi raised the issue of how the inadequacy of the current regime has also resulted in intervention by the judiciary which has created an entirely separate regime for content takedown in Sabu Mathew George v. Union of India & Ors, as well as Prajwala v. Union of India. Thus, a systematic approach is required to create a more comprehensive set of policies.

Divij Joshi also addressed the confidentiality requirement under Rule 16 of the blocking rules made under the framework of section 69A, which makes it impossible to ascertain the legitimacy or legality of the government action in any instance of blocking. While there is a need for some form of confidentiality to protect victims of cyber abuse, imposing broadly worded confidentiality clauses on private entities that have economic incentives to go overbroad in its application is problematic.

Reform and Alternative Models to Consider

According to Divij Joshi, there could be two approaches to content takedown mechanisms. The ultimate goal, or the goal to move towards would be the broader approach, which limits the concentration among platforms, and transfers that power to users to determine

what kind of content moderation frameworks should apply to them. A narrower approach, on the other hand would require a framework within the law that sets standards for intermediaries to follow.

The need for better accountability mechanisms was emphasised upon and aspects of Germany's Network Enforcement Act (NetzDG) were discussed as a possible model to emulate. NetzDG requires social network providers to produce reports on the handling of complaints about unlawful content on their platforms. Such a mechanism allows the public to scrutinise response to content takedown requests and thereby enables the formulation of better-informed policies. Similarly, NetzDG has also laid down different time frames for takedown of content depending on the nature of the content. In contrast, the intermediary guidelines rules have a blanket time frame for blocking of content which has now been decreased from 36 to 24 hours with no underlying rationale. Bhavna Jha argued that this is unfeasible as there cannot be one single time frame for takedown of different kinds of content. Content that is manifestly unlawful such as child pornography or extremist speech, should be filtered out automatically. However, this requires the surrounding structures that facilitate the implementation of such a framework. Divij Joshi referred to the Internet Watch Foundation's Child Sexual Abuse Imagery Database which has been hashed and shared with social media intermediaries to ease removal of child sexual abuse images. He further observed that such structures are necessary in implementing a framework that envisages the use of automated filtering to takedown unlawful content.

Post *Shreya Singhal*, the content takedown regime under section 79 is largely reliant on a court order or a government notification. Bhavna Jha suggested the incorporation of alternative regimes that empower people in content takedown. For instance, the content takedown system under New Zealand's Harmful Digital Communications Act, which prescribes a notice-to-notice takedown regime, was considered as an alternative, where the intermediary merely notifies the author of the content of the complaint. However, the intermediary is also empowered to take down content where such originator is unresponsive. Further, where the author of the content does submit a counter-notice refusing to consent to removal of the specific content, it can be challenged by the complainant in court. Thus, the legality of online content is adjudicated in court, which it was claimed was required in place of unilateral decision-making by the intermediary.

Differentiated Responsibility for Differentiated Content

The panel then discussed the scope for differentiated responsibility for different kinds of content. To this extent, archaic criminal laws such as obscenity and blasphemy laws, Bhavna Jha argued, should be re-conceptualised in order to place individual dignity and consent at the centre of content moderation. As an example, sexually explicit imagery was offered. Such content, in today's context, could be moderated if those images were being circulated without the consent of the persons in the image, rather than on the basis of outdated obscenity laws. Divij Joshi suggested that the legal framework needs to differentiate between the harm caused by different kinds of content. Without such

regulations, private entities would largely focus only on copyright moderation since that was the most profitable. This was illustrated in the discussion with the example of Content ID, a system developed to take down copyright infringements on Youtube. However, no such system existed for gender abuse or caste abuse online since there was no economic incentive to develop the same. Content moderation needed to aim at automated and expeditious removal of certain forms of content such as child sexual abuse imagery or extremist speech. On the other hand, automated takedown would not be as appropriate for copyright infringements. Therefore, there was a huge scope for differential treatment of content depending on what kind of harm it causes to users.

The final subject of discussion was the importance of accountability and transparency in keeping social media platforms in check. Various international frameworks such as Santa Clara Principles and the new Facebook Oversight Board were brought up by. It was noted that the principles had an immediate impact by outlining minimum levels of accountability and transparency that corporations need to maintain. Bhavna Jha also suggested the auditing of algorithms that set norms for content filtration and takedown. However, all mechanisms were to be accompanied with more overt and publicly accountable action that can only be done through the law.

Traceability

The third session was themed on the traceability obligation under Rule 3(5) of the Draft Intermediary Guidelines Rules.

Panelists

- Aditi Agrawal (Senior Research Associate, MediaNama)
- Anand Venkatanarayanan (Cybersecurity researcher)
- G S Madhusudan (Principal Scientist, IIT Madras)
- Shashank Mohan (Counsel, Software Freedom Law Centre)
- *Moderator*: Tanaya Rajwade (Policy Officer, CIS)

Whatsapp Traceability Case

The discussion on traceability and government intent was juxtaposed against the WhatsApp traceability case. Aditi Agrawal provided context for the proceedings that had taken place with respect to the case. The petition had originally been in the context of a cyber-bullying case. When the Tamil Nadu Police had been impleaded onto the case, they had mentioned the difficulties they faced in targeting fake news, and the perpetrators who spread disinformation. It was in this context that the scope of the petition was expanded to include law enforcement, and traceability came up in the discussion. After several hearings, the case had been transferred from the Madras High Court to the Supreme Court of India. The Central Government has repeatedly demanded traceability from platforms, with the Minister for Electronics and Information Technology, Shri Ravi Shankar Prasad having even made a statement to the effect that traceability would be a requirement, though how it could be implemented would be determined by the platform. The draft Rule 3(5) which requires intermediaries to enable the 'tracing out' of the 'originator' of information on its platforms will need to be considered in the context of both the Whatsapp traceability case, as well as the government's stance on the subject.

Implications for Social Media Platforms

For platforms such as Whatsapp, implementing changes to incorporate traceability had several implications. It would require them to break their end-to-end encryption. It would also mean that if such change was allowed in India, it would warrant changes in their service worldwide, which would then affect communications in more repressive regimes. Aditi also discussed the arguments made by various intermediaries, including Facebook Twitter, Whatsapp and YouTube. The arguments brought out the fact that different platforms work differently, and that already public platforms such as Twitter, or public posts on Facebook, would find it relatively easier to comply with a government order mandating traceability. However, it was argued in court that those platforms, specifically, messaging platforms like Whatsapp would find it near impossible to comply with a traceability obligation as a result of their end-to-end encryption. Additional arguments included the need to preserve freedom of speech and the right to privacy, as well as the

difficulties with making India-specific changes to account for the traceability requirement.

The IIT-M Proposal

What does the proposal say?

V. Kamakoti, a professor at IIT Madras, had been appointed as *amicus curiae* to the Madras High Court in the Whatsapp traceability case. He was specifically asked to come up with a proposal that enables the identification of originators of unlawful information without breaking end-to-end encryption, and ensuring that due process can be followed for 'traceability' requests. While there was some debate between Aditi Agrawal and G S Madhusudan regarding the contents of the proposal, the discussions that had taken place amongst the parties involved in the case, and what the judges of both the Madras High Court and the Supreme Court had stated on record, a brief summary of the IIT-M Proposal, is as follows:

The Court did not want end-to-end encryption to be broken and required the rule of law to be followed for such traceability requests. IIT Madras suggested a mechanism where Whatsapp encapsulates an identifier associated with the originator (the phone number) with each message created by this individual. This information is encrypted with a key such that only Whatsapp can decrypt this 'originator information.' A recipient of a message with unlawful content could alert law enforcement authorities of such content. These authorities could, after due diligence, issue a request to the intermediary, being Whatsapp in this case. Whatsapp may then carry out its own due diligence to decide the legitimacy of the request, decrypt the 'originator information' and reveal the identity of the originator of the message to the law enforcement agency. The proposed mechanism was aimed to (theoretically) uphold the due process requirement under law by shifting the control from the hands of the law enforcement agency to the intermediary, which has the prerogative to reject the request. With respect to the application of right to privacy and anonymity as recognized in the Puttaswamy judgment, IIT-M's position was that, in the case of WhatsApp, users voluntarily waive these rights when they use WhatsApp considering their numbers are visible and all messages they send are forwardable. On this aspect, the IIT-M proposal took the stand that WhatsApp should enable a feature that allows users to flag messages as 'not forwardable.'

Criticisms of the IIT-M proposal

One of the criticisms of the proposal brought forward by Shashank Mohan, was with respect to the possibility that such a process would concentrate discretionary power at the hands of the intermediaries, leading to conflicts between law enforcement and intermediaries. This could result in inefficiencies, and every matter being referred to the courts for determination. G S Madhusudhan's response in this regard was that this would be inevitable, since court systems are inherently adversarial, and therefore not meant to be efficient.

Another concern was regarding the implementation of the proposal. The discussion in this context revolved around whether the IIT-M proposal could even be implemented within its architecture in the first place. The claim made by Anand Venkatanarayanan was that the Signal protocol[10] embedded within Whatsapp architecture would mean that the theoretical proposal by IIT-Madras could not be incorporated. If incorporated, the end result would be an entirely new product altogether. Anand Venkatanarayanan suggested that the reason courts may have overlooked this issue, might have been due to the courts' own limited understanding of end-to-end encryption.

Concerns introduced by the draft Rule

Aditi Agrawal also mentioned that one of the arguments made during the Whatsapp hearings was regarding the intermediaries' obligation to assist law enforcement agencies under Section 69 of the IT Act, 2000. While Whatsapp had interpreted this to mean that they were to provide assistance 'to the extent possible', the government had interpreted it to mean that it could even require an intermediary to change its architecture. A possible side effect of introducing Rule (5), would be that law enforcement officials could allege that an intermediary was not cooperating in accordance with Section 69.

Concerns were also raised regarding the possibility that the power that would be created by the 'backdoor' that the traceability requirement seemed to mandate, would result in stifling opposition and dissent. Shashank Mohan discussed how this would also pave the way for state-backed agents of disinformation, or just any bad actor to take advantage of the backdoor. Concerns regarding digital attribution via traceability were raised by Aditi Agrawal, since there would still be the risk of misidentifying and wrongful prosecution, since traceability was not an absolute proof of who a sender could be. The *Puttaswamy* judgment had mandated that restrictions on privacy rights must be lawful, proportionate and for a legitimate state aim. Shashank Mohan also highlighted how the lack of any procedural safeguards meant that the Puttaswamy criteria was not met by Rule 3(5).

Discussions at Whatsapp's internal meetings suggesting encryption-at-the-edge methods to filter content, were also highlighted by Shashank Mohan to show how just the act of state pressure towards traceability have resulted in companies considering alternative mechanisms that would still have effects on the right to free speech. The growing coalescing of state demand to put an end to end-to-end encryption including by states such as the UK, US and Australia was brought up by Aditi Agrawal in context with the pushback by privacy advocates, Facebook and civil society.

---

[10] This article by Anand V. describes Signal protocol as, "*[...] where encryption keys used by a receiver and sender are continuously discarded and new keys are generated and used for the next message, once a previous message is sent and received successfully. This is described as a "Ratchet" (irreversible process). Since the intermediary can never see or store these encryption keys, it would not be possible for a government agency to demand the unencrypted messages between a sender and the receiver.*"
Anand Venkatanarayanan, 'Dr Kamakoti's solution for Whatsapp Traceability without breaking encryption is erroneous and not feasible' (*Medianama*, 13 August, 2019)
<https://www.medianama.com/2019/08/223-kamakoti-solution-for-traceability-whatsapp-encryption-madras-anand-venkatanarayanan/> accessed 17 June 2020

Less Restrictive Mechanisms

While the discussion on whether any less restrictive mechanisms could have been incorporated, instead of the traceability requirement, remained inconclusive, some suggestions and alternatives were discussed in this regard.

Another solution Shashank Mohan proposed was the use of metadata WhatsApp collects and provides to the law enforcement agencies on proper request along with an increase in law enforcement capacity, instead of looking at breaking end to end encryption. However, Anand Venkatanarayanan was of the opinion that the metadata which WhatsApp collects was very minimal and was insufficient evidence against any individual. He further asserted that it had been , in order to fill this void that the government in restricted circumstances had therefore considered hacking phones.

Conclusion

In conclusion, the panellists were in agreement that breaking end-to-end encryption was not a recommended path for governments to mandate. The panellists highlighted India's bad track record with protecting its databases, the profiling risks associated with pooling data across India's databases and the due process concerns if traceability is introduced as an obligation under law.

The future of Intermediary Liability in India

The final session was themed around rethinking intermediary liability in India, and the effect that the draft Intermediary Guidelines could have on the present understanding of Section 79 of the IT Act.

Panelists

- Alok Prasanna (Senior Resident Fellow, Vidhi Centre for Legal Policy)
- Sarvjeet Singh (Executive Director, Centre for Communication Governance)
- Tanya Sadana (Principal Associate, Ikigai Law)
- Udbhav Tiwari (Public Policy Advisor, Mozilla)
- *Moderator*: Gurshabad Grover (Research Manager, CIS)

Comparing historical and current intermediary liability contexts

The discussion commenced with each panellist making opening remarks regarding the significance of intermediary liability and the principle of safe harbour. Different arguments were put forth regarding the history behind intermediary liability in digital space and the purpose of a safe harbour provision. On one hand, Tanya Sadana argued that the purpose of safe harbour provision was to ensure that an intermediary is able to provide the best service to its users by exempting liability of intermediaries for user-generated content. In this context, it was argued that the draft intermediary guidelines inappropriately delegated state function to intermediaries by making them police user content. On the other hand, Alok Prasanna was of the opinion that the purpose of a safe harbour provision was to protect a nascent industry by ensuring that the cost of enforcing criminal laws is not placed upon intermediaries. Therefore, given how the industry is no longer at its nascent stage, and the power of intermediaries has significantly grown, it becomes necessary to examine questions beyond just the issues with the IT Act and the Draft Rules. While discussing the cost of enforcement, Sarvjeet Singh spoke about how smaller intermediaries must also be considered since they do not have the capacity to meet the obligations proposed under the draft rules. Udbhav Tiwari also discussed the different contexts within which US and Europe had evolved its intermediary liability regime, wherein the intent had been to grant the intermediary some freedom to regulate content in the manner in which they chose. This was compared to the Indian approach, which seemed to consider all intermediaries to be mere conduits. It was suggested that modern law on intermediary liability must answer questions around transparency and accountability, and set out the difference between accountability and liability. Lastly, it was emphasised that the ideal regulations surrounding intermediary liability would be that which ensures safety on the internet, while protecting speech to the extent possible.

Classifying Intermediaries under the IT Act

The discussion then moved on to deliberate upon the need for classification of intermediaries under the current legal regime. The panellists agreed on the need for distinguishing between intermediaries and to impose different obligations on different kinds of intermediaries. However, the panellists disagreed as to whether such a classification could be made under the intermediary guidelines itself. Sarvjeet Singh argued that no such classification can be made under delegated legislation where the parent statute makes no distinction between different intermediaries. Additionally, rules which provide for the active takedown of content by a distinct class of intermediaries would contradict the passive role of intermediaries as envisaged under section 79. On the other hand Udbhav Tiwari and Alok Prasanna were of the opinion that the definition of intermediaries allows for classifications to be done under the rules, even if section 79 does not explicitly say so. If one were to look at the development of administrative law in courts, rules tend to be struck down only if they directly contradict the parent legislation, or if the rules encompass a subject beyond the scope of the parent legislation. Courts have been cognizant of parliamentary constraints, and given a lot of leeway for parliament to delegate powers to the government. It was therefore argued that, as long as the categorization has a rational connection with the obligations imposed upon the intermediaries, such classification would be valid.

Basis of Classification of Intermediaries

The basis of such classification was also discussed. The panellists suggested that classification of intermediaries along various parameters, including size and function of the intermediary, its level of participation, the control it exercises, as well as the content dealt with by the intermediary. Alok Prasanna was of the opinion that no single criteria can be used to distinguish one intermediary from another. On this premise, he suggested as an alternative to be identifying a specified list of intermediaries (which might have to face a higher threshold of responsibility) within the law. Udbhav Tiwari added that such a framework must have due process and also incorporate guidelines and an appellate authority to prevent arbitrary inclusion and exclusion of intermediaries within the list. Udbhav Tiwari proposed the framework within the data protection bill as an example in this regard. E-commerce companies were also distinguished from other intermediaries, and it was debated whether the same criteria of liability would apply to them, given their level of participation in the selling of products. He further argued that e-commerce companies should not be able to claim safe harbour, given that the regulation of commerce had to be distinct from the regulation of speech.

Paradigm shift in functioning of Intermediaries

The panel then deliberated on a possible reconceptualization of section 79 on account of its strait-jacketed approach and underlying philosophy. Section 230 of the Communications Decency Act in the United States was considered as a rudimentary alternative. Tanya Sadana pointed out how a strict application of Section 79 would mean

that none of today's intermediaries (excluding internet service providers) would qualify for safe harbour under Section 79. Most intermediaries do intervene at the level of deciding the content that people see, even if the content is not modified. Chronological timelines have transitioned into algorithmically determined timelines that drive greater engagement. Alok Prasanna was of the opinion that, while it may be possible to read some nuance under Section 79 to account for these changes, at this current stage Indian laws seem to find it difficult to understand how to manage the element of subjectivity that has crept into the functioning of intermediaries. In this context, Udbhav Tiwari emphasised that an overhaul of Section 79 would require substantial efforts to be made, along the same lines as was done with the drafting of the Personal Data Protection Bill. The US example was used to show how the amendment of the section 230 exemption under the CDA, addressing content relating to online sex trafficking, had resulted in the unintended consequence of increased violence against sex workers.[11]

Differentiated Liability of Intermediaries and content creators

The discussion then shifted to the question of the need for differentiating the liability of intermediaries and content creators, and whether that was an additional reason to reconceptualize section 79. While the provision sets out how intermediaries can claim immunity from liability upon fulfillment of certain criteria, it does not explicitly lay down the consequences if an intermediary fails to meet the criteria. A paradigm with differentiated liability of intermediaries would pose practical difficulties in determining the extent of liability. However, Tanya Sadana suggested that principles such as abetment in the Indian Penal Code can possibly assist in the same. Finally, Alok Prasanna proposed that the question of distinct liability of intermediaries and the extent of the same can be determined by courts on a case-by-case basis as opposed to being over-determined by the law.

The session ended with panellists making closing remarks emphasising on due process and the duty of diligence in content-takedown and the need to incorporate principles of transparency and accountability. Panellists also debated the limits of state sovereignty in the times of social media, and suggested the need for looking towards the effective implementation of competition law principles as a remedy against existing issues of monopolies and network effects.

[11] See Katie Fiefer et al., 'Do Prostitution Advertisements Reduce Violence against Women? A Methodological Examination of Cunningham, DeAngelo and Tripp Findings' (2019) 4(3) DIGNITY: A JOURNAL ON SEXUAL EXPLOITATION AND VIOLENCE <https://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1167&context=dignity> accessed 17 June 2020