Comparison of the Manila Principles to Draft of The Information Technology [Intermediary Guidelines (Amendment) Rules], 2018

By **Akriti Bopanna** and **Gayatri Puthran**
Edited by **Elonnai Hickok** and **Torsha Sarkar**

September 30, 2019

Introduction

In December 2018, the Ministry of Electronics and Information Technology (MeitY) introduced amendments to the draft Information Technology [Intermediaries Guidelines (Amendment)] Rules, 2018 ["the 2018 Rules"]. The proposed changes ranged from asking intermediaries to proactively filter content using automated technology to prohibiting promotion of substances such as cigarettes and alcohol. In CIS's submission[1] to the Government, we highlighted our various concerns with the proposed rules. Building on the same, this paper aims to assess how the new draft rules measure up to the best practices on Intermediary Liability as prescribed in the Manila Principles. These principles were formulated in 2015 by a coalition of civil society groups and experts, including CIS, in order to establish best practice to guide policies pertaining to intermediary liability.

Depending on their function, intermediaries have a varying hand in hosting activism and discourse that are integral to a citizen's right to freedom of speech and expression. The Manila Principles are an attempt at articulating best practices that lead to the development of intermediary liability regimes which respect human rights.

Consequently, the paper examines the draft rules to assess their compatibility with the Manila Principles. It provides recommendations such that, where needed, the rules are aligned with the aforementioned principles. The assessment is done based on the insight into the rationale of the Manila Principles provided in its Background Paper.

[1] Grover, G., Hickok, E., Basu, A., Bopanna, A., & Sarkar, T. (2019, January 31). *Response to the Draft of The Information Technology [Intermediary Guidelines (Amendment) Rules] 2018*(Rep.). Retrieved https://cis-india.org/internet-governance/resources/Intermediary Liability Rules 2018.pdf

Principle 1: Intermediaries should be shielded by law from liability for third party content

Part (a) : Any rules governing intermediary liability must be provided by laws, which must be precise, clear, and accessible.

Constitutional validity

The draft Rules prohibits content which goes beyond constitutional limits on free speech and provisions. This is in direct contravention to key standards established by the Supreme Court of India.

The Rules under Section 79 of the Information Technology Act have been challenged before in *Shreya Singhal v Union of India*[2], in which the court held that any restrictions on content should only derive from the constitutional restrictions on free speech in Article 19(2) of:

a. Sovereignty and Integrity of India
b. Security of the State
c. Friendly relations with foreign States
d. Public order
e. Decency or morality
f. Contempt of Court
g. Defamation
h. Incitement to an offence

The terms of prohibited content in Rule 3(2) such as " harmful, harassing, pornographic, blasphemous, hateful, racially or ethnically objectionable, disparaging" go beyond what is permissible according to Article 19(2) of the Constitution.

In rule 3(2)(j), public health and safety is the criteria for the restriction which again, is not one of the constitutional restrictions on speech. Additionally, content that promotes cigarettes and other tobacco products, consumption of intoxicants including alcohol, ENDS and similar products is prohibited under the 2018 Rules save as approved by the Drugs and Cosmetics Act, 1940. However, the 1940 Act, which is the major governing legislation in this regard, only regulates the sale of nicotine gum containing up to 2gm of nicotine (as per

[2] Shreya Singhal. Union of India,AIR 2015 SC1523

Chapter IV of the Act), ENDS and like products. It does not cover the promotion of cigarettes, tobacco products or alcohol.[3]

The *Shreya Singhal* judgment clarified the "actual knowledge" needed for intermediaries to take down content under S.79(3)(b) is to be read as upon receiving a court order or order from an authorized government agency. The original Rule 3(4) of the Information Technology (Intermediaries guidelines) Rules, 2011 ["The 2011 Rules"] stood in contravention of this decision as it directed intermediaries to remove content upon their own knowledge of illegal information or upon being contacted by an affected party of the same, and prescribed a duration of thirty-six hours for the intermediary to take corrective action. In furtherance of the *Shreya Singhal* judgment, Rule 3(4) of the 2011 guidelines has been replaced by Rule 3(8) of the 2018 guidelines which permit takedown of content on the receipt of a relevant court order or government notification to that effect.

While this is welcome, Rule 3(9) of the 2018 Rules mandates intermediaries to employ automated technology to filter unlawful data. This delegation of legislative functions to private parties is unconstitutional because even such delegation has limitations and cannot be in the form of formulating new legislative rules.[4] Such a discernment of unlawfulness is the prerogative of the legislature and subsequently, the judiciary and not an executive arm such as the MeitY.

Rule 3(7) asks intermediaries to be incorporated as a company, have a physically registered office and appoint a nodal person for contact with law enforcement if either they have more than fifty lakh users or they are notified by the Government. The latter condition is not accompanied with any details on the reasoning behind the criteria of notifying an intermediary and consequently, leaving it open for the Government to exercise arbitrariness in making such a choice. Moreover,it is ultra-vires of S.79 of the IT Act since the provision merely concerns itself with the availment of safe harbor protection and instances that would make an intermediary not being able to avail such an immunity. By imposing rules as to the nature of the entity and manner of operation of an intermediary, the rule is in excess of the parent legislation. Similarly, Rule 3(10) asking for the reporting of cyber security incidents has no nexus to Intermediary Liability under S.79 and should instead be issued under S. 70B of the IT Act. Such ultra-vires provisions are legally void in nature.

On constitutionality, Principle 1.(a) of the Manila principles goes as far as to say that:

> *"Even when a law is constitutional, this does not necessarily mean that an intermediary should comply with it. If that law contravenes international human rights standards, and if the intermediary does not operate from that country or otherwise subject to its jurisdiction, then the intermediary is both legally and ethically*

[3] Grover, G., Hickok, E., Basu, A., Bopanna, A., & Sarkar, T. (2019, January 31). *Response to the Draft of The Information Technology [Intermediary Guidelines (Amendment) Rules] 2018*(Rep.). Available at: https://cis-india.org/internet-governance/resources/Intermediary Liability Rules 2018.pdf
[4] Ibid

justified in declining to enforce laws that would restrict the availability of its content within its borders"[5]

The International Covenant on Civil and Political Rights (ICCPR) under its A.19(3) states that any restrictions on free speech should be necessary and provided for in law in a manner that is clear, accessible and predictable.[6] The grounds for such restrictions should be of national security, public order, public health or morals and those which are needed to respect the rights or reputation of others.[7] India having ratified the ICCPR should have laws that adhere to these parameters and make the 2018 rules consistent with international human rights framework.

Prescribed by law

Any new rule needs to be within the ambit of the parent legislation under which it is being notified. Rule 3(9) extends beyond the scope of S. 79 of the IT Act by placing new responsibilities on the intermediaries such as asking for messages to be traceable, proactive filtering of content using technology etc. S.79 states that intermediaries are not liable for the information they host if they do not:

"*(i) initiate the transmission,*

(ii) select the receiver of the transmission, and

(iii) select or modify the information contained in the transmission;"[8]

If any more obligations need to be assigned to intermediaries then it should flow from a modification to the IT Act or creation of a new one, not as an executive notification under the Act.

Lack of definition

One of the central issues of the 2018 Rules is the term, 'unlawful content', and its usage. This is because it is neither defined in the 2018 rules or the Information Technology (IT) Act. Other Indian statutes including the Indian Penal Code make certain actions illegal or unlawful but do not define what unlawful content could mean. The gamut of things seen as unlawful as

---

[5] Manila Principles Background Paper 1.0. (2015, July 09). Available at <https://www.eff.org/document/manila-principles-background-paper-10>

[6] Grover, G., Hickok, E., Basu, A., Bopanna, A., & Sarkar, T. (2019, January 31). *Response to the Draft of The Information Technology [Intermediary Guidelines (Amendment) Rules] 2018*(Rep.). Available at https://cis-india.org/internet-governance/resources/Intermediary Liability Rules 2018.pdf

[7] International Covenant on Civil and Political Rights - available at <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>

[8] Section 79 of the Information Technology Act, 2000

per the 2018 Rules, without it being stated as illegal in any legislation, goes against the spirit of this principle.

Secondly, Rule 3(2) asks intermediaries to inform users that they are prohibited from hosting or engaging with 'unlawful' content that includes content that is "disparaging", "racially, ethnically or otherwise objectionable", "relating or encouraging money laundering or gambling". These terms are largely undefined. Specifically in Rule 3(2)(j), the words "threaten" and "public health or safety" are not explained or defined in any of the laws referenced by the Rules. The subsequent provision on "threatening" critical infrastructure also leaves the word undefined.

Rule 3(5) is ambiguous due to lack of a definition of the words; "assistance" as well as "protective or cyber security".

Lastly, no definition is provided for terms used in Rule 3(9); "automated technology" and no standards laid down for the phrase; "appropriate mechanisms/control".

Part (b) : Intermediaries should be immune from liability for third-party content in circumstances where they have not been involved in modifying that content

In this matter, the drafters of the principles referred to the 2011 Joint Declaration on Freedom of Expression and the Internet which states:

> *"No one who simply provides technical Internet services such as providing access, or searching for, or transmission or caching of information, should be liable for content generated by others, which is disseminated using those services, as long as they do not specifically intervene in that content or refuse to obey a court order to remove that content, where they have the capacity to do so ("mere conduit principle").*[9]

This aspect of the principle flows from the previous one, to emphasize that intermediaries should be given maximum protection by law. The rationale being that when the content on platforms is created by the users and not by the intermediaries themselves, the liability for the illegality of any such content must not fall on them.

According to Rule 3, apart from just modifying, intermediaries cannot knowingly host, publish or transmit illegal content. In addition to holding intermediaries liable for actions beyond modification, this rule does not take into account that on many platforms, intermediaries do not have express control on the data that is being transmitted or published, and thus places liability on companies for aspects they cannot control.

---

[9] Manila Principles Background Paper 1.0. (2015, July 09). Available at <https://www.eff.org/document/manila-principles-background-paper-10>

Under the Manila Principles, Principle 1.(b) does not provide a form of blanket immunity since intermediaries are liable for any modifications they do undertake. Modification can either be content-related or technological and while any changes to content that contravene existing laws will negate the immunity, technological changes that do not relate to the actual content itself should not be penalized. This would include , opting for a higher standard of encryption to display content of a website.

The 2018 Rules itself in 3(11) permit for such technological changes stating;

> *"The intermediary shall not knowingly deploy or install or modify the technical configuration of computer resource or become party to any such act which may change or has the potential to change the normal course of operation of the computer resource than what it is supposed to perform thereby circumventing any law for the time being in force. Provided that the intermediary may develop, produce, distribute or employ technological means for the sole purpose of performing the acts of securing the computer resource and information contained therein."*[10]

However, the ambit of Rule 3(2) is excessively wide with undefined terms that go beyond the law, which makes liability for content on intermediaries much larger than appropriate, and envisioned under the Manila Principles. Importantly, the 2018 Rules also contrast the approach in the Principles in that the latter establishes a general principle of non-liability of intermediaries unless they have modified third party content. Under the Rules, intermediaries are liable for modifying (and hosting, publishing or transmitting) third party content unless the modification is technological in nature and justifiable under security reasons.

Part (c) : Intermediaries must not be held liable for failing to restrict lawful content.

While the provisions in the 2018 Rules concern unlawful content, the lack of definitions for the kinds of content to be filtered by the intermediaries essentially seeps into regulating content that may well be legitimate, and lawful. Without legal clarity on parameters such as "ethnically objectionable", "disparaging" and so on, intermediaries will have to resort to using their judgment on the same, resulting in potential restriction of lawful content.

Further, the Principles note the intersection of data protection and intermediary liability within this rule, in that legislation on the former can ask intermediaries such as search engines to restrict access to lawful content. In the European Union, for instance, where the

[10] Comments/suggestions invited on Draft of "The Information Technology [Intermediary Guidelines(Amendment) Rules] 2018", Ministry of Electronics and Information Technology, 2018,<http://meity.gov.in/content/comments-suggestions-invited-draft-%E2%80%9C-information-technology-intermediary-guidelines>

right to be de-indexed from search engines has arisen out of the right to be forgotten.[11] This aspect can give rise to ambiguity for a search engine between making all legal content available and citizen's right to erasure of their data.[12] So far in the Indian context, S.27 of The Personal Data Protection Bill (if passed) qualifies the right to be forgotten to a restriction or prevention of ongoing disclosure of their personal data. However, this is a right that is available to citizens subject to an Adjudicating Officer determining that the restriction of data overrides the right to free speech and right to information of citizens.[13] Thus, there is another layer of decision making and responsibility between the citizen and intermediary, reducing the decision making ability of the latter to modify access to content.

A compromise arrived at, in some jurisdictions, is the implementation of a notice and notice regime. With potentially large varieties of content being classified as unlawful under the rule, a notice and notice system would reduce the liability of intermediaries for the content where the content itself is not manifestly illegal. Their responsibility would be to forward any order to the end user, challenging the legality of such content. This could be a form of striking a balance between the State's need to regulate potential harms emanating from absolute free speech and the one to keep the online domain a space of individual and collective expression.

Part (d) : Intermediaries must never be made strictly liable for hosting unlawful third-party content, nor should they ever be required to monitor content proactively as part of an intermediary liability regime

Rule 3(9) is directly in contravention of this principle by asking intermediaries to proactively filter unlawful content. In doing so, it is unclear if intermediaries would be held liable for failing to proactively filter the content. This principle also calls for intermediaries to only be held liable for not complying with a takedown order from the court or a government agency, and not for hosting it in the first place. Rule 3(3) includes hosting and transmission, rather than just modification, of illegal content.

In a 2011 study carried out by CIS, we saw that a lack of clear guidelines on content takedown can lead to a chilling effect on speech with intermediaries removing more content than needed in order to be on the safe side of law.[14] This fear could potentially be amplified with the rule to filter content using technology. Such a rule negates the spirit of the Shreya

---

[11] Article 17 of the General Data Protection Regulation. Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679>

[12] Manila Principles Background Paper 1.0. (2015, July 09). Available at <https://www.eff.org/document/manila-principles-background-paper-10>

[13] Section 27 of Draft Personal Data Protection Bill, 2018. Available at <https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf>

[14] Dara, R. (2011). Intermediary Liability in India: Chilling Effects on Free Expression on the Internet. Retrieved from https://cis-india.org/internet-governance/chilling-effects-on-free-expression-on-internet

Singhal judgment, according more power to an intermediary than was permitted by reserving the right for such content removal decisions to be executed by a Court.

Recommendations

1. Terms like hateful, racially, ethnically objectionable and disparaging in Rule 3(2)(b) should be deleted as they are not defined under Indian law.
2. The term unlawful in 3(2)(b) and 3(9) should be deleted.
3. Rule 3(2)(j) should be deleted for going beyond the limits prescribed in Article 19(2)
4. Rule 3(3) must be amended to only include modification as a ground on which intermediaries can be held liable rather than hosting or transmission since intermediaries often do not initially have the knowledge of or control over content being posted onto their websites.
5. Rule 3(7) and 3(10) should be deleted.
6. Rule 3(9) must be struck down since it tasks intermediaries to decide what unlawful content is. Instead, a rule similar to Article 15 of Directive 2000/31/EC must be established which establishes a notice and takedown approach, requiring compliance from intermediaries but states that it is not their responsibility to monitor content on its platform.[15]

Principle 2 - Content must not be required to be restricted without an order by a judicial authority

Part (a) : Intermediaries must not be required to restrict content unless an order has been issued by an independent and impartial judicial authority that has determined that the material at issue is unlawful

This principle advocates for removal of content by intermediaries only as per the direction of a competent judicial authority. It acknowledges that this places a high burden on the judiciary in having to assess all content takedown requests.

Rule 3(9) places the responsibility of filtering and restricting unlawful content on intermediaries, through the use of automated technology. This is clearly in opposition to this Principle 2.(a) which aims to establish judicial orders as the only viable mode of a content restriction. The dangers of the same are highlighted before in this paper. Not all relevant intermediaries have the financial or institutional resources to employ automated technology to begin with, and further one trained well enough that can filter content as per the 2018

[15] Directive on Electronic Commerce, available at <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32000L0031>

Rules demands. Youtube, a major tech incumbent, has invested a total of 100 million dollars on ContentID[16], an automated tool responsible for finding copyright violations in the market, and even then, it has proved to be inaccurate.[17] Some might have to withdraw their services altogether if the burden of compliance is too high. Eventually, it contributes to stifling of innovation with only the bigger and better funded entities being able to align their operations with the Rules.

The Principles also suggest a notice and notice system to reduce the burden on the intermediary and accord more responsibility to the users. An expedited judicial structure for the sole purpose of tackling content takedown matters would also be a significant step in resolving these matters efficiently.

Part (b) : Content of the orders of restriction

The following are the requirements of content restriction orders as prescribed by the M*anila Principles* in their background paper*:*

1. *Provide a determination that the content is unlawful in the jurisdiction.*
2. *Indicate the Internet identifier and description of the unlawful content.*
3. *Provide evidence sufficient to document the legal basis of the order.*
4. *Where applicable, indicate the time period for which the content should be restricted.*

The gamut of things seen as unlawful as per the 2018 Rules, without it being stated as illegal in any legislation, goes against the spirit of this principle. The Rules simply state that a court order by a judicial authority or a notification by a government body will be required to affect content removal, without undertaking what the order or notification should constitute in terms of its content. Thus, the dearth of such guidelines results in the absence of a standard for the kind of direction that intermediaries may receive. The specificity and precision of the order or notification may be compromised and this affects the accuracy with which content takedown is carried out. Further, the time periods for content removal have been indicated but not how long the content will be made inaccessible for.

Part (c) : Any liability imposed on an intermediary must be proportionate and directly correlated to the intermediary's wrongful behavior in failing to appropriately comply with the content restriction order

The 2018 Rules do not mention any penalties for non-compliance and do not hold the intermediaries criminally liable. They do, however, place excessive measures of compliance

[16] Sawers, Paul. 2018. "YouTube: We've invested $100 million in Content ID and paid over $3 billion to rightsholders". Available at <https://venturebeat.com/2018/11/07/youtube-weve-invested-100-million-in-content-id-and-paid-over-3-billion-to-rightsholders/>

[17] King, Ashley. 2019. "Youtube is addressing Major issues with Vague Contentid Claims". Available at <https://www.digitalmusicnews.com/2019/03/21/youtube-contentid-changes>

on the intermediaries. The time period allowed to intermediaries to comply with the directions provided by courts or governmental agencies has been reduced from 36 hours to 24 hours, thereby making compliance more difficult and increasing the propensity to be in breach of the requirements for availing safe harbour. In the case that the government requests information, previously according to the 2011 guidelines, no time period was prescribed within which the same would have to be made available, but according to the 2018 Rules, 72 hours have been mentioned.

These time restrictions, having been made more stringent in the 2018 Rules, increase the likeliness of intermediaries losing their safe harbour protection. Many intermediaries, especially start-ups, may not have the resources or the funds to be able to take down content or compile information within the given time period. The infeasibility of these provisions, and the increased risk of losing their immunity, may lead to intermediaries being over-cautious with content on their platform and result in takedown of contentious content without court or government orders, as a precautionary measure to evade liability.[18]

Part (d) : Intermediaries must not be liable for non-compliance with any order that does not comply with this principle

Recommendations

1. Rule 3(9) should be struck down, owing to the resultant over-censorship issues that come with allowing intermediaries to use automated technology
2. An expedited judicial process should be put in place to specifically address content takedown requests. This can be more efficient by virtue of bypassing full court proceedings, decreasing the burden on judicial institutions, and enabling faster resolution of content takedown disputes while also redressing the rightholder more efficiently.
3. There should be a mandated list of details in takedown orders, in line with the Manila Principles.
4. Rule 3(5) must be struck down since the time period mentioned applies to all intermediaries. Instead, a more nuanced timeline considering the different kind of intermediary and their functions should be put in place.
5. Instead of mandating a traceability requirement in rule 3(5), section 69(1) and 69B of the IT Act, Section 91 and 92 of the Criminal Procedure Code and Section 5 of the Telegraph Act and 419A Rules of the allied rules should be referenced and relied on.

[18] *Intermediaries, users and the law – Analysing intermediary liability and the IT Rules*(Rep.). (n.d.). Available at: <https://sflc.in/sites/default/files/wp-content/uploads/2012/07/eBook-IT-Rules.pdf>

Principle 3 - Requests for restrictions of content must be clear, be unambiguous, and follow due process

Part (a) : Intermediaries must not be required to substantively evaluate the legality of third-party content.

The legal implications of Rule 3(9) have been discussed earlier in the paper, but additionally, the efficacy of such automated technology is questionable at best. This technology has not evolved enough to understand the intricacies of human discourse, such as sarcasm and satire, and its use could lead to censorship of legitimate information or speech. Existing research shows that such technologies have a significant error rate and there needs to be human intervention to review what content is being filtered out[19]. For instance, a critique of the European Commission's proposal to use filtering technology is that automated technology would not be able to "distinguish between Syrian war footage used by ISIS and the same footage used by human rights advocates"[20]. Filtering technologies can only detect if the file matches prohibited content and not "whether a particular use of an identified file is an infringement in light of the context within which the media was being used."[21] The aftermath of mandated use of automated filtering technologies can potentially see a spike in content takedowns and account suspensions, with asymmetrical effects on groups such as journalists, activists and dissidents.[22]

The use of inexact automated technology to detect illegal content threatens the rights of users since decision making on content is then based on unclear and unaccountable rationale. Coupled with the lack of appeal options to such removals, the sanctity of due process is contravened. Further, rule 3(8) does not require intermediaries to give a notice to user content uploaders when their content is being taken down thus leading to an "invisible" form of censorship that violates the norms of due process.Additionally, there is no process of appeal within this provision or in Rule 3(4) where the optional termination of account by

---

[19] MacCarthy, Mark, and I. *AI-Driven Content Moderation Can Never Be Perfect*. CIO, 26 Oct. 2018, available at <www.cio.com/article/3316562/artificial-intelligence-driven-content-moderation-can-never-be-perfect.html.>

[20] Daphne, Keller. *Problems with Filters in the European Commission's Platforms Proposal*. Center for Internet and Society, 5 Oct. 2017, Available at <cyberlaw.stanford.edu/blog/2017/10/problems-filters-european-commissions-platforms-proposal>

[21] Engstrom, Evan, and Nick Feamster. *The Limits of Filtering: A Look at the Functionality & Shortcomings of Content Detection Tools*. Engine, 2017, *The Limits of Filtering: A Look at the Functionality & Shortcomings of Content Detection Tools*,. Available at <static1.squarespace.com/static/571681753c44d835a440c8b5/t/58d058712994ca536bbfa47a/1490049138881/FilteringPaperWebsite.pdf>

[22]"Intermediary Liability 2.0 - A Shifting Paradigm." *SFLC.in*, Mar. 2019, sflc.in/intermediary-liability-20-shifting-paradigm.

intermediaries can be effectuated without informing the user or giving them a mechanism of appeal.

Part (b) : A content restriction request pertaining to unlawful content must, at a minimum, contain the following:

1. The legal basis for the assertion that the content is unlawful.
2. The Internet identifier and description of the allegedly unlawful content.
3. The consideration provided to limitations, exceptions, and defenses available to the user content provider.
4. Contact details of the issuing party or their agent, unless this is prohibited by law.
5. Evidence sufficient to document legal standing to issue the request.
6. A declaration of good faith that the information provided is accurate.

As it stands today, content restriction orders for non-copyright infringement grounds do not have a legislatively defined or judicially approved format. The 2018 Rules lack information on the nature and contents of the takedown order, as elaborated in the previous section. It is vital to have details such as the specific URL of the content because clearly defined orders will decrease the onus of intermediaries to monitor content, in general. It will also ensure that intermediaries do not take down more content than needed when trying to comply with the order. Most importantly, the more defined the order is, the better position an intermediary will be in to question orders that do not comply with the letter of the law.

Under Rule 3(5), any government agency can seek any information or assistance related to the "*security of the state or cyber security, investigation or detection or prosecution or prevention of offence(s), protective or cyber security*" and related matters. Such assistance requested is not defined in the Rules and thus, could take place in the form of a content restriction notice. For the same, a lawful order is needed which the government agency can seek, merely by sending in a request in writing or electronically with the purpose for the information or assistance.

Firstly, a lawful order does not necessarily mean a court order which implies that content can be restricted without a court order. Significantly, it is any government agency which can request for such assistance, without there being any requirement for them to be legally authorized ones. The latter part of the Rule specifically states that intermediaries shall enable tracing of the originator of information on their platforms only by legally authorized agencies. This distinction created in the same rule gives rise to concern since a government agency need not have legal standing to issue a content takedown request.

This provision should be deleted from the rules and existing laws under the IT Act such as S.69 (1) and 69B of the IT Act should be relied on, to access information and assistance

The 2018 Rules mandate a grievance officer to whom such complaints can be referred. Currently, Rule 75 of Copyright Rules read along with Section 51 of the Copyright Act is the

sole reference for content of notices wherein it asks a copyright holder to have certain information in their notice takedown request to an intermediary which is a mere conduit.[23]

Part (c) : Content restriction requests pertaining to an intermediary's content restriction policies must, at the minimum, contain the following:

1. The reasons why the content at issue is in breach of the intermediary's content restriction policies.
2. The Internet identifier and description of the alleged violation of the content restriction policies.
3. Contact details of the issuing party or their agent, unless this is prohibited by law.
4. A declaration of good faith that the information provided is accurate.

There are several restrictions under Rule 3(2) that are not defined in the law, as discussed earlier in the paper. In combination with the continued absence of clear guidelines on what information should be mentioned in content restriction requests, it lends an ambiguous environment for intermediaries to operate in.

Part (d) : Intermediaries who host content may be required by law to respond to content restriction requests pertaining to unlawful content by either forwarding lawful and compliant requests to the user content provider, or by notifying the complainant of the reason it is not possible to do so ('notice and notice'). Intermediaries should not be required to ensure they have the capacity to identify users.

The 2018 Rules do not have any provisions that require a notice and notice procedure.

Part (e) : When forwarding the request, the intermediary must provide a clear and accessible explanation of the user content provider's rights, including in all cases where the intermediary is compelled by law to restrict the content a description of any available counter-notice or appeal mechanisms

The Principles recommend that the intermediary's role be as a notifier between the aggrieved party seeking the removal and the end user who has uploaded the said content. However, the 2018 Rules expand the intermediary's role with the obligation to proactively filter content. There is no appeal process under any situations of restrictions under the rule. As a result, there is no onus on explaining the reasons for takedown or any opportunity to

[23] Joshi, D. Indian Intermediary Liability Regime. Available at <https://cis-india.org/internet-governance/files/indian-intermediary-liability-regime>

contest this. The user should have a forum to provide for a defence to the alleged illegality whether its through low cost arbitration as suggested in the Principles or through any other judicial recourse.

Part (f) : If intermediaries restrict content hosted by them on the basis of a content restriction request, they must comply with Principle VI on transparency and accountability below.

Transparency principles have not yet been laid down in the rules.

Part (g) : Abusive or bad faith content restriction requests should be penalized.

There is no specific penalty or disincentive in the guidelines so far and hence no legal provision to bar frivolous requests.

Recommendations

1. A list of legally authorized government agencies who can issue a notification for content removal must be determined and made publicly available.
2. Rule 3(5) should be amended to include a requirement to show the information/assistance is based on the grounds of a) the security of the state b) investigation, detection, prosecution or prevention of an offence
3. In rule 3(5), the ground of “protective or cyber security or matters [...]or incidental thereto” must be deleted since it is undefined.
4. A rule should be added to the guidelines listing the details that a content restriction request must contain.
5. A rule in the 2018 Rules must establish the requirement of a notice and notice procedure.
6. Users should have a forum to defend the content they have uploaded against the alleged illegality.
7. Transparency and accountability measures must be built into the 2018 Rules and bad faith restriction requests should be penalized.

Principle 4: Laws and content restriction orders and practices must comply with the tests of necessity and proportionality

Part (a) : Any restriction of content should be limited to the specific content at issue

There is no provision in the 2018 Rules to address the issue of specificity required in content restriction orders, they do not require requests to be narrow or limited to specific types of content.

Part (b) : When restricting content, the least restrictive technical means must be adopted.

There is no provision that limits the content that can be taken down. Rule 3(9) makes it incumbent upon intermediaries to remove illegal content through technological methods without restricting or regulating this power or framing any guidelines about the extent of it.

Part (c) : If content is restricted because it is unlawful in a particular geographical region, and if the intermediary offers a geographically variegated service, then the geographical scope of the content restriction must be so limited

The 2018 Rules do not address matters of extraterritoriality when it comes to terms of enforcement. Though, they make it compulsory for entities having more than fifty lakhs users to register as a company under the Companies Act, 1956 apart from having a registered office. This can be seen as an attempt to adhere to regulating content and entities not based in India, thus largely expanding the scope of jurisdiction by Indian government agencies and law enforcement authorities.

Part (d) : If content is restricted owing to its unlawfulness for a limited duration, the restriction must not last beyond this duration, and the restriction order must be reviewed periodically to ensure it remains valid.

The 2018 Rules do not assign a time duration for the suspension of the content. In effect, content can be taken down indefinitely without a mechanism for review or reinstatement. The lack of provisions that require: specificity in requesting content to be removed, a time

frame for content removals and in which territories can inadvertently lead to legal and lawful data being taken down.

Recommendations

1. The 2018 Rules should stress on the requirement of specificity in terms of the content that is being restricted and the geographical area over which this restriction is valid.
2. A rule is needed to clearly establish the technical means to be adopted to enforce a restriction and the time duration over which the restriction takes place.

Principle 5 - Laws and content restriction policies and practices must respect due process

Part (a) : Before any content is restricted on the basis of an order or a request, the intermediary and the user content provider must be provided an effective right to be heard except in exceptional circumstances, in which case a post facto review of the order and its implementation must take place as soon as practicable.

This part requires the following:

1. The content uploader should be present at court proceedings deliberating content restriction. The party that uploads the content has a right to be heard.
2. The above does not apply to exceptional circumstances, but there must be an ex post facto review in those circumstances.
3. These exceptional circumstances include but are not limited to (1) child pornography, (2) direct and public incitement to commit genocide, (3) advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence, (4) and incitement to terrorism.
4. The three part test for content restriction policy is (1) unambiguous law (2) pursuance of legitimate purpose (3) respect for principles of necessity and proportionality

Rule 3(8), when addressing matters related to the court process only mention that intermediaries will have 24 hours within which they will have to comply with a content restriction order. These guidelines do not delve into the details of who should be present during the court proceedings or the existence of any rights for the content uploading party, let alone what these rights may constitute. The 2018 Rules do not acknowledge the rights of the content uploader in regular circumstances, by extension, they thus neglect both defining any exceptional circumstances under which such rights may be temporarily deposed and the requirement of an ex post facto review.

The third point categorizes these exceptional circumstances, which are not limited to the ones listed. The 2018 Rules do not make any such categories or any attempt to define "exceptional circumstances".

Rule 3(4), especially, does not incorporate due process mechanisms by allowing intermediaries the right to terminate access of users and their rights without informing them of the same.

Lastly, this principle lays down the standards for content restriction policies. Rule 3(4) merely mention that intermediaries must inform their users monthly about the need to comply with their policies and regulations while Rule 3(2) lays down a host of content restrictions. While these restrictions are problematic and drafted amgiously, the 2018 rules do provide some standards such as content cannot be blasphemous. However, these have the effect of impinging on fundamental rights.

Part (b) : Any law regulating intermediaries must provide both user content providers and intermediaries the right of appeal against content restriction orders

There is no specific provision in the 2018 Rules that provides for intermediaries' and content providers' right of appeal against content restriction orders.

Part (c) : Intermediaries should provide user content providers with mechanisms to review decisions to restrict content in violation of the intermediary's content restriction policies

The Rules provide for a grievance redressal mechanism for parties affected by content on the intermediary's platform, but there is no reciprocal equivalent mechanism for content uploaders to appeal against an intermediary's decision to take down their content.

Part (d) : In case a user content provider wins an appeal under (b) or review under (c) against the restriction of content, intermediaries should reinstate the content

The 2018 Rules do not make it incumbent upon intermediaries to have a mechanism in place where user content uploaders can challenge content takedowns and ask for a review, and thus, by extension there is no mechanism in place that requires content to be reinstated if it were found that the takedown was wrong.

Part (e) : An intermediary should not disclose personally identifiable information about a user without an order by a judicial authority. An intermediary liability regime must not require an intermediary to disclose any personally identifiable user information without an order by a judicial authority

The Principles elaborate that:

> "Governments should not legally require intermediaries to disclose personally identifiable information as part of an intermediary liability regime without a judicial order. To this extent, Governments should not hold an intermediary liable for failing to disclose personal data of users without such an order" [24].

However, according to rule 3(5) of the guidelines *"The intermediary shall enable tracing out of such originator of information on its platform as may be required by government agencies who are legally authorised"*. Intermediaries under Rule 3(5) need to provide the government with information and assistance for cases "concerning security of the State or cyber security; or investigation or detection or prosecution or prevention of offence(s); protective or cyber security and matters connected with or incidental thereto." Thus, in the aforementioned circumstances, the intermediary has to share personally identifiable information to governmental agencies without a judicial order. The lack of clarity on what a lawful order is, provides more loopholes to gain citizen's information without legal backing.

Part (f) : When drafting and enforcing their content restriction policies, intermediaries should respect human rights. Likewise, governments have an obligation to ensure that intermediaries' content restriction policies respect human rights

There is no reference or provisions in the 2018 Rules that require content restriction policies to protect human rights. Lack of transparency can pose as a danger to minority groups or opinions that are more susceptible to their freedom of expression being violated, which constitutes as a massive human rights violation. In a white paper by Electronic Frontier Foundation[25], it is evidenced that automated technology used by intermediary platforms have the effect of silencing marginalized voices. Here also, the lack of transparency on the part of these platforms regarding the error rate of their automated technology keeps human

---

[24] Manila Principles Background Paper 1.0. (2015, July 09). Availble at <https://www.eff.org/document/manila-principles-background-paper-10>

[25] York, J. C. (2019, June 03). Caught in the Net: The Impact of "Extremist" Speech Regulations on Human Rights Content. Available at https://www.eff.org/wp/caught-net-impact-extremist-speech-regulations-human-rights-content#Content

rights violations in the dark. In lieu of a suggestion in a report by the United Nations Special Rapporteur[26], human rights impact assessments can be regularly carried out to verify whether company policies and terms of service are mindful of human rights and whether they underpin due process.

Recommendations

1. The 2018 Rules should include provisions for the rights of intermediaries and user content providers, including "right to be heard" and right to post facto review of content removals.
2. The 2018 Rules should include a provision regarding rights of intermediaries and user content providers to challenge content takedown orders.
3. Intermediaries must be transparent about the content restriction directions they receive from the government. Google maintains such a report of government requests for content takedowns[27]. A similar apparatus should be put in place for content taken down privately by intermediaries for being non compliant with their terms and regulations, since this form of censorship is private rather than public[28]. Additionally, there must be a mechanism for user content providers to challenge these takedowns. Thus the 2018 Rules must first guide intermediaries to establish and update transparency reports and consequently institute portals for users to challenge takedowns.
4. A provision needs to be inserted into the 2018 Rules to enable reinstatement of content that has been wrongfully taken down. To enable this, the users must be notified of the reason their content was taken down. If, upon review, the takedown is found to be wrongful, there must be an accessible mechanism for users to demand reinstatement of this content. The new provision must describe such a mechanism and prescribe a timeframe within which the reinstatement must take place.
5. In rule 3(5) the last sentence from "The intermediary shall enable tracing out of such originator [...] required by government agencies who are legally authorised." should be deleted since it infringes on the privacy of citizens.
6. The 2018 Rules must have provisions that stress the importance of transparency, mandate human rights impact assessments and require mandatory disclosure about content that has been taken down through automated technology.

---

[26] Content Regulation in the Digital Age(Rep.). (2018, February).Available at <https://www.ohchr.org/Documents/Issues/Opinion/ContentRegulation/APC.pdf>

[27] Google Transparency Report. (n.d.). Available at <https://transparencyreport.google.com/government-removals/by-country?hl=en>

[28] Singh, S., & Bankston, K. (2018, October). The Transparency Reporting Toolkit: Content Takedown Reporting(Rep.). Available at <https://d1y8sb8igg2f8e.cloudfront.net/documents/The_Transparency_Reporting_Toolkit_Content_Takedown_Reporting_2018-10-24_125414_1.pdf>

Principle VI. Transparency and accountability must be built into laws and content restriction policies and practices

Part (a) : Governments must publish all legislation, policy, decisions and other forms of regulation relevant to intermediary liability online in a timely fashion and in accessible formats

Currently, the IT Act and its rules, along with the comments and counter-comments relating to intermediary liability are available on the website of the Ministry of Electronics and Information Technology. However, relevant judicial decisions and content takedown orders sent by the Government are not made available by them. The former are accessible through academic databases which are not open access and require a significant fee. On the other hand, numbers around content restriction orders are found if and only when intermediaries publish their transparency reports and these are often not standardised. Moreover, civil society has only been able to access lists of restricted content by filing Right to Information requests.

Part (b): Governments must not use extra-judicial measures to restrict content. This includes collateral pressures to force changes in terms of service, to promote or enforce so-called "voluntary" practices and to secure agreements in restraint of trade or in restraint of public dissemination of content

The 2018 Rules compel intermediaries to publish their rules and regulations, privacy policy and user agreement as per rule 3(1) and inform its users that non-compliance with the same gives the intermediary the rights to remove the infringing content and terminate the usage or access rights of the content provider as per rule 3(4). The rules and regulations are required to prohibit content that falls under many vaguely described terms such as content that is “grossly harmful, harassing, blasphemous, defamatory, obscene [...] hateful, or racially, ethnically objectionable, disparaging”as mentioned under rule 3(2).

Part (c): Intermediaries should publish their content restriction policies online, in clear language and accessible formats, and keep them updated as they evolve, and notify users of changes when applicable

While Rule 3(1) does direct intermediaries to publish their usage agreements and rules and regulations, it neglects informing them to keep it updated and to notify users about any key

changes when applicable. The result of this could be users being unaware of the changing policies and rationale for their content being taken down.

Part (d): Governments must publish transparency reports that provide specific information about all content orders and requests issued by them to intermediaries

India has the Right to Information Act through which such data can be requested, but the government itself does not have any requirement to publish transparency reports on its own. The 2018 Rules have no provisions that require the same, either from the government or from the intermediaries.

Part (e): Intermediaries should publish transparency reports that provide specific information about all content restrictions taken by the intermediary, including actions taken on government requests, court orders, private complainant requests, and enforcement of content restriction policies

The 2018 Rules do not have any such requirements for intermediaries. While some intermediaries, like Google or Facebook, do publish some amount of information on the same, it is still not as detailed as required by the Manila Principles.

Part (f): Where content has been restricted on a product or service of the intermediary that allows it to display a notice when an attempt to access that content is made, the intermediary must display a clear notice that explains what content has been restricted and the reason for doing so.

Nothing in the 2018 Rules requires intermediaries to display messages about the reason of inaccessibility of any content on their website.

Part (g): Governments, intermediaries and civil society should work together to develop and maintain independent, transparent, and impartial oversight mechanisms to ensure the accountability of the content restriction policies and practices

No such mechanism has been established in the 2018 Rules.

Part (h): Intermediary liability frameworks and legislation should require regular, systematic review of rules and guidelines to ensure that they are up to date, effective, and not overly burdensome. Such periodic review should incorporate mechanisms for collection of evidence about their implementation and impact, and also make provision for an independent review of their costs, demonstrable benefits and impact on human rights

There is no general obligation to review the rules or laws.

Recommendations

1. Rule 3(4) should be amended to direct the intermediaries to notify users about changes in the regulations as well.
2. The guidelines should introduce a rule to direct governments and intermediaries to publish transparency reports which are in line with the Data Protection bill and eventual, legislation.
3. Such guidelines should include a provision that directs intermediaries to display messages explaining the reason for specific content takedowns.

Conclusion

The 2018 Rules are largely unaligned with the Manila Principles and require a significant amount of change. Several of its provisions have the potential to infringe upon the right to free speech. Further, the rules fail to put in place any mechanisms to promote transparency and accountability. They delegate undefined powers to governmental authorities to censor and control the content hosted and transmitted on intermediary platforms. Further, they necessitate the use of automated technology, which has questionable effectivity, to filter content. Several amendments are needed before the guidelines are consistent with a model that envisions intermediaries to be a mode of facilitation of expression, as opposed to one that engages in discerning morality and unlawfulness.

At the same time, there is a need to look at the revision of the Manila Principles, keeping in mind the contemporary regulatory circumstances. Developments such as the proliferation of extremist propaganda on social media websites and harassment of minorities, for instance, necessitate quicker responses from intermediaries. Jurisdictions are introducing, or experimenting with regulation that mandate such interventions, like the European Union with its regulation on 'preventing the dissemination of terrorist content online'.[29] Article 4 of the same, imposes a one hour time framework for hosting service providers to remove the terrorist content or disable access to it.[30] Thus, future research in this area can focus on analyzing the continued relevance of the Manila Principles.

---

[29] Available at <https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-preventing-terrorist-content-online-regulation-640_en.pdf>

[30] Ibid