MATTHIAS C. KETTEMANN, JOSEFA FRANCKE, CHRISTINA DINAR AND LENA HINRICHS (EDS.)

Platform://Democracy

Research Report Asia-Pacific

PLATFORM://DEMOCRACY

Platform://Democracy

Perspectives on Platform Power, Public Values and the Potential of Social Media Councils: Research Report Asia-Pacific

edited by Matthias C. Kettemann, Josefa Francke, Christina Dinar and Lena Hinrichs

LEIBNIZ INSTITUTE FOR MEDIA RESEARCH | HANS-BREDOW-INSTITUT, HAMBURG, GERMANY
HUMBOLDT INSTITUTE FOR INTERNET AND SOCIETY, BERLIN, GERMANY

LEIBNIZ INSTITUTE FOR MEDIA RESEARCH HANS-BREDOW-INSTITUT

ALEXANDER VON HUMBOLDT INSTITUTE FOR INTERNET AND SOCIETY

STIFTUNG MERCATOR

Contributors

Name(s)	Affiliation
Setu Bandh Upadhyay	Technology Policy and Digital Rights Analyst
Bayan Koshravi	Mehralborz Higher Institute of Education, Tehran, Iran
Lucinda Nelson	Queensland University of Technology, Brisbane, Australia
Prateek Waghre	Internet Freedom Foundation, New Delhi, India

Table of Contents

Introduction to the Asia-Pacific Research Clinic

Setu Bandh Upadhyay
TECHNOLOGY POLICY AND DIGITAL RIGHTS ANALYST

The Asia-Pacific research clinic aimed at examining the normative infrastructure needed to establish better rulemaking, rule-enforcing, and rule-adjudication structures in hybrid communication spaces, primarily online. Specifically, the clinic sought to explore the potential of platform councils as a solution to address legitimacy deficits in private and hybrid orders of platforms, which are characterized by idiosyncratic normative logics, vocabularies, and algorithmic arrangements. The clinic also examined the decentralization of power and alignment of public values with private orders, while exploring various governance models for social media platforms.

The clinic focused on the Asia-Pacific region and brought together a diverse set of perspectives. Discussions delved into nuanced topics such as the impact of social media councils when dealing with different languages, fact-checking, Web 3.0, communication infrastructure, local platforms, value-sensitive design, legitimacy of self-regulation, and transparency. The clinic's approach encouraged the fellows to engage in short, but deep diving into specific regional challenges instead of going for broad overviews. In this way, each short paper was able to contribute meaningfully to the debate, and created some appetite for more deep dives into regional avenues.

The clinic was composed of experts in social media governance, journalism, social justice, human rights, law, technology policy, digital media research, and platform governance. The fellows represented the Asia-Pacific region, including Iran, India, Sri Lanka, Singapore, Hong Kong, Taiwan, and Australia. Additionally, the participants had expertise in other jurisdictions such as China, East Asia, and South Asia.

Themes and Outputs

Governments can encourage or 'nudge' the adoption of private ordering through platform councils as a way to improve the platform ecosystems. Civil society, scholars, and activists are advised to closely monitor the level of market concentration and the rate of innovation in platform ecosystems. This is to ensure that the benefits of platform councils outweigh any risks associated with it. For example, platform councils may lead to increased competition and innovation, but it could also result in the exclusion of certain groups or the exploitation of vulnerable individuals. Policymakers may consider exempting smaller platforms and open knowledge community platforms from major platform regulations and compliance requirements that are meant for larger platforms.

Regarding specific platforms with regional challenges, one study explored the role of social media as the main channel of communication in schools and how value-oriented and democratic design can be leveraged to improve social media governance.

Another study focused on regional case studies in an effort to study how cultures and norms of acceptable behavior are shaped on social media platforms. Using feminist theory, the study showed that Social Media Councils can publicly expose gender inequality in content moderation by calling attention to instances where content moderation policies or practices unfairly target or harm women or other gender minorities and that they play a crucial role in shaping rules relating to gender equality through their interpretations of vague policies in specific cases.

One study criticized the executive-appointed social media council in India with a focus on their formation, legitimacy, capacity, scale, and expertise. The study highlighted that an executive-appointed

and controlled council is not desirable due to concerns over low levels of operational transparency and vagueness of remit across multiple bodies and intermediaries.

The Impact of Private Ordering on Platform Competition

Setu Bandh Upadhyay
TECHNOLOGY POLICY AND DIGITAL RIGHTS ANALYST

Key Findings

This paper explores how private ordering is linked to competition. The paper uses theoretical analysis to establish linkages due to a lack of regional examples. The paper concludes with recommendations that talk about nudging by regulators/government for platforms to adopt private ordering, encouraging a deeper exploration of the relationship between private ordering and competition, and suggests private ordering as a solution to facilitate a competitive platform ecosystem.

Introduction

Social media platforms like Meta, Twitter, and YouTube have become vital to the global information landscape. How they manage user-generated content has significant implications for freedom of expression, privacy, and public discourse. This is where private ordering comes into the picture. Platforms like Facebook and Instagram already have the Oversight Board to regulate their content policies and decisions. And Twitter also had a Trust & Safety council of independent experts that was dissolved by Elon Musk. Spotify, a music platform with a social feature, also has a safety council. While this has provided a degree of democratic voice in the governance and regulation of platforms, it has also raised concerns about censorship, bias, and the influence of dominant players in the industry.[1]

The advantage of private ordering is that it allows for a more flexible approach to regulation. Businesses and industries are often better equipped to understand their operations' unique challenges and requirements and can take proactive steps to self-regulate. For example, platforms can respond to emerging trends in online harassment, hate speech, and misinformation by implementing new policies and tools to address these issues. This can result in a more effective and efficient response than government regulation, which often moves at a slower pace. However, private ordering by social media platforms can also have drawbacks, such as a lack of transparency, accountability, and consistency. The platforms may not always be transparent about their content moderation policies and may enforce these policies in inconsistent or discriminatory ways. In some cases, platforms may act to limit free speech, suppress dissent, or otherwise undermine democratic values and human rights.[2]

Private ordering can also create barriers to entry for new competitors, limit consumer choice, and entrench dominant players. The impact of private ordering on platform competition depends on the specific design of the private agreements and the broader regulatory and market context in which they

---

[1] Newton, C. (2022). To build trust, platforms should try a little democracy. Retrieved February 14, 2023, from Platformer at https://www.platformer.news/p/to-build-trust-platforms-should-try.

[2] Kettemann and Fertmann (2022). Platform-proofing Democracy—Social Media Councils as Tools to Increase the Public Accountability of Online Platforms. HIIG. Retrieved February 13, 2023, from https://www.hiig.de/publication/platform-proofing-democracy-social-media-councils-as-tools-to-increase-the-public-accountability-of-online-platforms/.

operate. As platforms move towards private ordering, it is crucial to understand this trend's impact on competition and the overall regulatory ecosystem in the context of platforms.

If platforms have a competitive advantage when they use platform councils, it may suggest that private ordering can potentially lead to market concentration by players enacting platform councils resulting in an overall reduction in competition. On the other hand, if platforms do not have a competitive advantage when they use platform councils, it may suggest that private ordering can promote competition and innovation. In either case, private ordering can play a role in promoting democratic control over markets to redistribute ownership and control of communication infrastructure by providing a voice on governance to people who are not only driven by profits.

Private Ordering and Competition

Empirical studies suggest that competition among platforms often responds in unanticipated and sometimes ambiguous market positions.[3] This is clearer when it comes to platforms that work on user-generated content – i.e., social media platforms. The research reveals that consumers' content preferences are governed by network effects, potentially indicating a correlation towards their intention to have a voice in content and platform governance.[4]

Industry standards set by private ordering can improve the efficiency and effectiveness of operations and increase trust between the industry and consumers. While platform councils can provide a flexible and adaptive approach to content moderation challenges, they can also create barriers to the entry of new players, community-generated or hosted platforms, and customer choice in the way of additional financial burden or increased internal compliance in jurisdictions where they face no regulatory burden.

Wikipedia has been a great example of providing a democratic and collective voice in the governance of a significant platform, now known as the "Wiki-Governance" model.[5] In 2022, The Indian Ministry of Electronics and Information Technology requested that the Wikimedia Foundation furnish details on their response to the defacement of Wikipedia pages relating to two individuals named Arshdeep Singh - an Indian cricketer and an Indian footballer.[6] This request was made in accordance with the Information Technology Rules, 2021, specifically sections 3(1)(b) and 3(1)(d), which prohibit intermediaries from hosting, storing, or publishing illegal information that poses a threat to the sovereignty and integrity of India, among other things.[7] Subsequently, the affected pages were placed under "semi-protection" to limit editing access to trusted users. The entire incident, including the adaptability Wikimedia could rely on using their moderators, demonstrates the benefits of private ordering, industry norms, and user control over the platform, which also seems to have no adverse effect on the competition.

Private ordering can have a significant impact on competition and innovation. On the one hand, private ordering can encourage competition by promoting uniformity in business practices and allowing for a level playing field. On the other hand, private ordering can also limit competition by creating new business entry barriers and stifling innovation. It can lead to an unfair competitive edge for dominant

[3] Cennamo, C., & Santalo, J. (2013). Platform competition: Strategic trade-offs in platform markets. Strategic management journal, 34(11), 1331-1350.

[4] Zhang, K., & Sarvary, M. (2011). Social media competition: Differentiation with user generated content. Marketing Science, 47, 48.

[5] Dove, E. S., Joly, Y., & Knoppers, B. M. (2012). Power to the people: a wiki-governance model for biobanks. Genome Biology, 13, 1-8.

[6] Agrawal, A. (2022). No, India didn't 'summon' Wikipedia over Arshdeep edits. It asked for information within 24 hours. Newslaundry. https://www.newslaundry.com/2022/09/06/no-india-didnt-summon-wikipedia-over-arshdeep-edits-it-asked-for-information-within-24-hours.

[7] PRS Legislative Research. (2022). The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. PRS Legislative Research. https://prsindia.org/billtrack/the-information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021.

players in an industry. When a few large companies control a large share of the market, much like the social media landscape, they may be able to use private ordering as another tool to entrench their position and limit the entry of new players.

This can lead to market concentration and anti-competitive behavior. In such cases, private ordering can stifle innovation and limit the ability of smaller, less established platforms to enter the digital sphere and sustain. This ends up reducing the number of choices available to consumers. One famous example of undemocratic private ordering which directly led to stifling competition has been the Motion Picture Association. The association used the film ratings system as an unreasonable restraint on trade to smaller studios and independent films.[8] This comparison with social media platforms makes sense because similar to MPA, there are few prominent players making decisions on what content may be viewed by users and not.

Moreover, if Platform Councils are not transparent and open to only participants who are driven by profit (e.g., the company appointed majority), it can lead to a lack of oversight and accountability, which can again result in anticompetitive behavior in the form of collusion and harm to consumers. When used responsibly and with oversight, Platform Councils can lead to more efficient, democratic, and effective regulation.[9]

Measuring the impact of private ordering on competition and innovation can be challenging, as it can be difficult to determine the precise effects of private ordering on competition policy in the digital age. Digital markets are have many crucial variables which can influence competition, such as multisided markets, powerful network effects, economies of scope/scale, large amounts of user data, disruptive innovations, integrations, and switching costs.[10] For these reasons, analyzing competition in digital markets and assessing whether changes are needed to existing competition policy frameworks with respect to private ordering needs in depth study of said features.

However, it is possible to assess the impact of private ordering on competitiveness by examining the level of market concentration, the rate of innovation, and the level of consumer choice. Although private ordering by social media companies can have positive effects on competition and innovation, it is important to carefully monitor the process to ensure that it is fair, transparent, and open to all participants. This will help to promote healthy competition, foster innovation, and ultimately benefit consumers.

---

[8] Kilburn, C. E. (2013). An Offer You Can't Refuse: A Sherman Act Antitrust Examination of the Motion Picture Association of America and the Use of the Ratings System as an Unreasonable Restraint on Trade. UMKC L. Rev., 82, 255.

[9] Klonick, K. (2020). The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression (SSRN Scholarly Paper No. 3639234). https://papers.ssrn.com/abstract=3639234.

[10] OECD (2022). OECD Handbook on Competition Policy in the Digital Age. https://www.oecd.org/daf/competition-policy-in-the-digital-age.

Recommendations

This piece is only able to cover selected aspects of private ordering and competition and ends with a call on competition scholars to analyze the impact of private ordering on innovation and competitiveness on platforms. As discussed above, private ordering offers several benefits, including quick adaptation, increased innovation, and personalization. However, on close inspection, it is revealed that private ordering also carries risks, including lack of oversight, anti-competitive behavior, and market concentration.

Therefore, it is important to monitor, explore and study the relationship private ordering has on specific industries, including platform businesses. The findings have important implications for stakeholders. Policymakers must consider the impact of government intervention on competition and innovation and the potential consequences of relying too heavily on private ordering.

The following key recommendations are proposed:

1. Platform councils and other forms of private ordering improve platform ecosystems while ensuring a fair, democratic, and competitive environment.
2. Civil society, scholars, and activists should pay careful attention to the potential of private ordering to increase or decrease the level of market concentration and the rate of innovation to ensure that the benefits of private ordering are greater than the risks.
3. Policymakers in Asia, particularly in places where technology policy and platform regulation are in nascent stages, may want to take inspiration from European Union's Digital Services Act. Smaller platforms and open knowledge community platforms should be exempt from major platform regulations and compliances.

Conclusion

Private ordering has the potential to significantly impact innovation and competitiveness both positively and negatively. It is, therefore, vital to carefully consider the benefits, risks, and challenges that come with private ordering in comparison to government intervention. By taking a balanced approach, policymakers can promote a healthy business environment that supports competition and innovation.

It is important for governments, civil society, activists, and other stakeholders to engage with social media platforms to promote a healthy balance between private ordering and public oversight. Platform councils may help deal with improving transparency, promoting accountability, and encouraging the development of best practices and industry standards. By working together through platform councils, regulators, platforms, users, and other stakeholders can help to promote a healthy and vibrant online environment that supports freedom of expression, privacy, and public discourse while also giving breathing room for a competitive environment and alternative platforms

Which role can Social Media Councils play in educational contexts? The case of the Shaad Platform in Iran

Bayan Koshravi
MEHRALBORZ HIGHER EDUCATION INSTITUTE, TEHRAN, IRAN

Abstract

I build on the findings of a case study to illustrate how Social Media Councils (SMCs) can have an effective role in an educational environment under conditions of high government interventions. Even though the Shaad platform was almost entirely regulated by government agencies, I argue that SMCs could play a significant role in the enforcement of rule of law and appeal mechanisms. I also argue that, in similar contexts, the proper design of SMCs must consider inclusion as the primary criteria for ensuring that enforcement and appeal decisions reflect democratic values such as diversity and participation. Lessons learned from this case could be useful in forming more inclusive SMCs.

Case Background: Shaad as a social media platform

On the eve of the Covid-19 lockdown and school closures in Iran, the Iranian government introduced Shaad as a communication and educational social network that was to provide the platform for teaching in public schools across the country for the next two years. This platform was a substitution for all teaching, communication and administrative tools that previously were available in person. It was developed by Iran's largest telecom operator on the basis of a general-purpose Iranian social network called Rubica. Shaad was offered free of charge and users incurred no cost for accessing content or data usage. Three roles were defined for users: students, teachers and headmasters. Nearly all Iranian students were automatically registered on the network, and it is estimated that about 70% of Iranian students used it daily. Despite the increased burden on parents to be involved in their children's education, no monitoring or supervisory role was created for parents on the platform.

While much of the content moderation services were provided by third parties (mainly teachers), the Ministry of Education of Iran made all defining decisions about how the system should be used in schools and what types of educational content can be provided on the platform. Teachers' primary responsibility was to moderate the content posted by students and to ensure that content used for educational purposes met the standards set by the ministry. There are also reports that suggest that certain groups of inspectors were set up in every province to monitor whether schools used Shaad. This issue was important to the government since it wanted to be sure that all students have access to educational content while schools were closed.

Establishing a social media council in this context could have a huge impact on the operation and further development of the Shaad platform. At the very least, there would have been ample opportunity for involving parents, teachers and other third parties like educational technologists in the regulatory process. There are parents-teachers associations in approximately every school in the country which meet regularly to discuss issues related to schools and the education of children. No arguments have been made against replicating this system of parents-teachers associations (or a modified version of it) on the platform. It seems that the opportunity simply escaped the minds of policy makers and parents

were far too unfamiliar with the platform and too concerned about the consequences of the Covid-19 pandemic to demand involvement.

Even though such councils were never formed, Shaad could benefit from the use of an SMC in an educational context. Some issues that an effective educational SMC can solve include safety and privacy issues, barriers in communication on the platform between different actors who should be involved in the educational process of students, and the presentation of inappropriate content.

These governing issues require a multifaceted approach that goes beyond government regulations. While regulations can set a standard for online behavior and safety, they are often too broad and generalized to provide effective solutions to the specific challenges faced by educational social platforms. An effective social media council is essential in providing tailored solutions to above-mentioned issues by taking into consideration the unique needs and values of the platform's users. Therefore, the diversity of expertise and ideas of SMC members will provide a holistic approach to governance that ensures the safety and well-being of users while promoting productive, engaging and meaningful educational experiences. Furthermore, an effective SMC has the flexibility to adapt and respond to changing trends and emerging challenges as they happen.

Case Analysis and Discussion

Since Shaad was limited to an educational context and was developed and operated wholly inside of the country, there were no political concerns about the flow of information that might challenge the government's authority. There was also no legal framework for the operation of such platforms. Activities on Shaad were only subject to broader guidelines concerning the educational system and online protection of children. Much of the government concerns were directed at dealing with online bullying, ensuring access to age-appropriate content and adherence to educational standards. None of these topics are politically controversial as they are nearly universally accepted. Therefore, a possible role of SMCs could be targeted towards applying and enforcing these rules and values on the platform. In this way, SMCs can be useful to address the problem of content governance. This role is now mostly limited to teachers whereas there is a high potential of using the capacity of other stakeholders.

However, this does not suggest that the design and operation of such councils would be an easy task. Iran is a vast country with diverse cultures and a complex ethnic composition. Thus, interpretation and adherence to regulation require local adaptation in ways that cannot be predetermined by central administration. Therefore, SMCs in similar contexts should be based on promoting multistakeholderism within the rule of law at a local level. Since multistakeholderism is often aimed at ensuring the inclusion and equal participation of stakeholders in the collective decision-making process, in this section, I discuss the principles that must underpin the design of such efforts in similar contexts.

Identity Management and User Training are Critical for Ensuring Inclusion in SMCs

Inclusion in the digital society is important to ensure the fairness of procedures through which goods or services are allocated[1] especially to marginalized groups and in order to enable them to voice their concerns and empower them to participate more fully in wider society.[2]

Digital inclusion can be defined as "effective participation of individuals and communities in all aspects of knowledge-oriented society and economic-oriented society by providing access to technology, by excluding availability barriers and enhancing the capability of communities to take ICT benefits"[3]. Scholars have identified several factors that contribute to inclusion, including access, digital literacy and digital skills[4].

Access is generally understood to be the precondition for participation in the digital society, even though it is by no means enough.[5] Identity management (i.e. a framework of policies and technologies to ensure that the right users have the appropriate access to technology resources) is an essential tool in ensuring 1) the participation of parents in distributed regulation of the platform, and 2) the availability of appropriate access for each user (e.g. teachers, students, parents, inspectors) according to their role. In the Shaad platform only students were recognized as users and in many cases siblings of different ages used the same device. For example, there are reports that older siblings used the platform to contact students from other schools (and often the opposite sex) because of the absence of a mechanism to ensure the logging out from their siblings' account who used the same device to access the Shaad. Moreover, parents had no means of ensuring appropriate use of the platform and occasionally had to use it to contact other parents or teachers as well. This also weakened the regulatory process because it could not be established who had used the device if an inappropriate message was sent. An SMC could play an effective arbitration role (in cases of misconduct) and push further development of the platform towards better identity management and inclusion of parents.

In addition, parents, teachers and other third parties should be aware of the capabilities and features of the platform, for example the type of content that can be uploaded and different options that one can use to communicate with other users. Otherwise, not everyone will be in a position to have an equal role in

---

[1] e.g. Azmi, A., Ang, Y. D., & Talib, S. A. (2016). Trust and justice in the adoption of a welfare e-payment system. *Transforming Government: People, Process and Policy*, *10*(3), 391-410.

Martin, A., & Taylor, L. (2021). Exclusion and inclusion in identification: Regulation, displacement and data justice. *Information Technology for Development*, *27*(1), 50-66.

Masiero, S., & Arvidsson, V. (2021). Degenerative outcomes of digital identity platforms for development. *Information Systems Journal*, *31*(6), 903-928.

[2] Maier, S., & Nair-Reichert, U. (2007). Empowering women through ICT-based business initiatives: An overview of best practices in e-commerce/e-retailing projects. *Information Technologies & International Development*, 4(2), 43-60;

Hassanin, L. (2008, September). Egyptian women artisans: ICTs are not the entry to modern markets. In *IFIP International Conference on Human Choice and Computers* (pp. 179-190). Springer, Boston, MA.

[3] European Commission. (2007). *European i2010 initiative on e-Inclusion: "To be part of the information society"*. Brussels: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2007) 694 final.

[4] Madon, S., Reinhard, N., Roode, D., & Walsham, G. (2009). Digital inclusion projects in developing countries: Processes of institutionalization. *Information Technology for Development*, *15*(2), 95-107.

Van Dijk, J.A.G.M. (2005). The deepening divide: Inequality in the information society, Sage Publications, Inc

Heeks, R. (2022). Digital inequality beyond the digital divide: conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, *28*(4), 688-704.

[5] Armenta, Á., Serrano, A., Cabrera, M., & Conte, R. (2012). The new digital divide: the confluence of broadband penetration, sustainable development, technology adoption and community participation. *Information Technology for Development*, 18(4), 345-353.

Friederici, N., Ojanperä, S., & Graham, M. (2017). The Impact of Connectivity in Africa: Grand Visions and the Mirage of Inclusive Digital Development, *Electronic Journal of Information Systems in Developing Countries*, 79(2), 1-20.

the decisions made. Without a diversity-oriented approach towards SMCs, the marginalized or powerless groups for example those from minority ethnic groups will find little or no opportunity to develop skills, exercise their rights to preserve their values and culture or play an autonomous role in content moderation on the platform. If not properly configured, SMCs may conceal actual power structures[6] (e.g. certain powerful ethnic groups or individuals who own or control the platform can influence and bias the content presented on the platform) and fail to achieve their objective of bringing democratic values to the governance of social networks. On the other hand, inclusion of marginal groups can enhance their sense of ownership and belonging to a community, inside the platform as well as outside. For instance, students and parents from a minority group may contribute actively to the platform by sharing their work, insights and this can motivate them to continue engaging with the platform and contribute more. This feeling of responsibility in the platform's success and taking active steps to maintain the platform's quality will contribute positively to the platform's growth and development.

Furthermore, in order to ensure that all stakeholders are aware of the opportunities for involvement in the governance of social media platforms, SMCs must meet the highest standards of transparency regarding their own operations. This could be exemplified by the encouragement of content production and participation by marginal groups, to which we now tend.

SMCs Must Go Further to Promote Inclusion through Encouraging Content Production by Marginal Groups

If the content and information on the platform are produced and used only in ways decided by dominant groups, marginal groups will not be adequately represented on these platforms and very likely will play no role in SMCs. Therefore, in educational contexts, SMCs must not limit their role to enforcing guidelines. Rather, they should actively encourage the creation and curation of content produced by marginal groups in order to ensure that they remain visible and can have a voice when critical decisions are to be made. This also ensures the display of diversity of values which can help maintain an open mind in discussing specific cases.

Ideally, Different Roles Defined in the Educational Context Must be Represented in the SMC.

The composition of members of SMCs in educational contexts must represent the various roles defined in this context. Therefore, it is necessary that teachers, parents and officials have representation and can play an active role in the functioning of such a council. Whether or not these roles will be equal cannot be determined without a careful consideration of broader legal context related to educational platform and distribution of accountability. Students can also be represented in these councils, although they are unlikely to have any voting rights because they are not mature enough. Additionally, individuals with experience in educational technology, and experts in software development, data security, privacy and other technical fields who can provide insight and advice to the council in platform design, features and best practices can also play an active role in SMCs. Additionally, leaders from local or national educational organizations, advocacy groups and associations who can provide guidance on policy and legal issues related to the educational platform should be represented in SMCs.

It is only by adhering to these principles that SMCs may contribute to the promotion of democratic values in the educational context as well as in society at large.

---

[6] Kettemann M.C. (ed.). (2022). How Platforms Respond to Human Rights Conflicts Online. Best Practices in Weighing Rights and Obligations in Hybrid Online Orders, Hamburg: Verlag Hans-Bredow-Institut.

Social Media Councils and Gender (In)Equality: An Analysis of Decisions by Meta's Oversight Board

Lucinda Nelson
QUEENSLAND UNIVERSITY OF TECHNOLOGY, BRISBANE, AUSTRALIA

Abstract

This short paper examines the potential role of social media councils (SMCs) in addressing gender inequality through an analysis of three relevant decisions of Meta's Oversight Board, as a current example of a platform-specific SMC. This analysis reveals the important roles the Oversight Board plays in publicly exposing gender inequality and improving transparency in Meta's content moderation, and the features that allow the Oversight Board to perform these roles, being: opportunity for public comment; sufficient independence from Meta; the ability to ask Meta questions about its policies and practices; publicly available reasoning; access to independent research; and an external set of rules to guide decision-making. This case study also highlights the Oversight Board's inability to directly effect changes to platform policies and enforcement practices, and the challenges this presents when Meta's policies and practices are discriminatory. Finally, this paper calls for further research into: whether the Oversight Board, or a different model of SMC, should be empowered to make binding decisions about platform policies and enforcement practices; and the potential for SMCs to address other axes of inequality, such as race and sexuality.

Context

Social media platforms have become an important space for public expression, discussion and deliberation. To date, platforms have mostly taken an undemocratic approach to content moderation, making decisions that significantly impact users' expression without public consultation or explanation. Women and gender diverse people are disproportionately impacted by content moderation policies and enforcement practices. Their ability to express themselves online is directly limited by discriminatory policies and practices,[1] and by a lack of moderation of harmful content that seeks to silence them.[2] In this context, SMCs have emerged as a potential mechanism for the development of better approaches to content moderation, including in relation to gender inequality. Meta's Oversight Board is just one example of a SMC, but it reveals interesting insights which can be applied in different contexts and in the development of new SMCs.

---

[1] Salty. 2021. Algorithmic Bias Report. https://saltyworld.net/product/exclusive-report-censorship-of-marginalized-communities-on-instagram-2021-pdf-download/.

[2] Khan, Irene. 2021. Promotion and Protection of the Right to Freedom of Opinion and Expression (No A/76/258, United Nations). https://documents-ddsny.un.org/doc/UNDOC/GEN/N21/212/16/PDF/N2121216.pdf?OpenElement.

Case study 1: Breast cancer symptoms and nudity[3]

This case involved the removal of an Instagram post containing photos of bare breasts with breast cancer symptoms, under the *Adult Nudity and Sexual Activity policy*.[4] Uncovered female nipples are generally prohibited under the *Adult Nudity and Sexual Activity policy* but are permitted if posted for 'educational or medical purposes', including breast cancer education. After this case was selected by the Oversight Board, Meta (at the time, Facebook) acknowledged the removal was a mistake, restored the post and urged the Oversight Board to decline the case on the basis that the mistake had been corrected. The Oversight Board rejected this argument, heard the case, and overturned the original decision to remove the post.

This case highlights the value of the Oversight Board in publicly exposing gender inequality in Meta's policies and enforcement practices. Content moderation is largely 'black box'.[5] Women and gender diverse people report being disproportionately targeted by content moderation removals,[6] but it is difficult to prove these claims without evidence of social media platforms' internal policies and practices.[7] By proceeding with the case, even after Meta had acknowledged the error and restored the post, the Oversight Board drew **public attention to gender inequalities** in automated content moderation. It found that Meta's reliance on automated enforcement of the *Adult Nudity and Sexual Activity policy* is likely to have a disproportionate impact on women because of the different treatment of 'male' and 'female' nipples. The case also provided an **opportunity for public comments** on the topic, which included arguments that Meta's nudity policies discriminate against women. Importantly, this decision could only be made because the Oversight Board was **sufficiently independent** from Meta to reject the proposal to decline the case.

Case study 2: India sexual harassment video[8]

This case involved the removal of a video posted on Instagram depicting a Dalit woman in India being sexually assaulted, under Meta's *Adult Sexual Exploitation policy*.[9] The removal was flagged for review internally after an employee learnt about it on Instagram. Meta then restored the post, with a warning screen, under a newsworthiness allowance. The newsworthiness allowance is broad and rarely applied, and involves balancing 'the public interest' and the potential for harm, without clear criteria. The Oversight Board upheld Meta's eventual decision to restore the post to Instagram. This decision was selected for analysis as sexual assault disproportionately affects women and is a key site of gender inequality.

---

[3] Oversight Board. 2020. 'Breast cancer symptoms and nudity.' https://www.oversightboard.com/decision/IG-7THR3SI1/.

[4] Meta. 'Adult nudity and sexual activity.' Accessed February 23, 2023. https://transparency.fb.com/en-gb/policies/community-standards/adult-nudity-sexual-activity/.

[5] Gillespie, Tarleton. 2018. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media. New Haven: Yale University Press.

[6] West, Sarah Myers. 2018. 'Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms.' New Media & Society 20(11): 4366-4384; Salty. 2021. Algorithmic Bias Report. https://saltyworld.net/product/exclusive-report-censorship-of-marginalized-communities-on-instagram-2021-pdf-download/.

[7] Cotter, Kelley. 2021. '"Shadowbanning Is Not a Thing": Black Box Gaslighting and the Power to Independently Know and Credibly Critique Algorithms' Information, Communication & Society: 1-18.

[8] Oversight Board. 2020. 'India sexual harassment video.' https://www.oversightboard.com/decision/IG-KFLY3526/.

[9] Meta. 'Adult sexual exploitation.' Accessed February 23, 2023. https://transparency.fb.com/en-gb/policies/community-standards/sexual-exploitation-adults/.

This decision reflects the value of the Oversight Board as a body for making complex, policy-shaping determinations more transparently. Traditionally, large social media platforms have taken a formal equality approach to content moderation, in part due to the difficulty of considering context at scale.[10] In this case, the Oversight Board was able to take a more substantive equality approach, considering a range of contextual factors relevant to both the public interest value of the content and its potential to cause harm, including the particular marginalisation of Dalit women in India. This contextual understanding was supported by **independent research** commissioned for this case. Importantly for **transparency**, the case was opened for public comment, and the Oversight Board's **reasoning is publicly available** for people to consider and critique in a way that decisions made solely by Meta are not. While transparency alone does not equal accountability, it is a prerequisite.[11]

Where the existing rules and exceptions are not clearly defined, as in this case, the Oversight Board plays a significant role in shaping platform rules through its precedential decisions. In these cases, the composition of the Oversight Board could have a critical impact on the decisions made, as different people are likely to have different perspectives on the public interest value of a piece of content. Members holding explicit or implicit sexist views, for example, would be likely to interpret any ambiguities in platform policies in sexist ways. The Oversight Board's reference to international human rights law, as a pre-determined, **external set of rules**, is useful to mitigate the impact of member perspectives. For the purposes of advancing gender equality, it is also important that board members understand and are **genuinely committed to addressing gender inequality.**

Case study 3: Gender identity and nudity[12]

These two cases (bundled) involved the removal of two Instagram posts consisting of photos of a transgender person and a non-binary person, both bare chested, under the *Sexual Solicitation* community standard.[13] The Oversight Board overturned Meta's decisions to remove the posts in question on the basis that the nipples were not 'female' and were therefore permitted under an exception.

This decision demonstrates the role of the Oversight Board in improving transparency by **exposing inconsistencies** between Meta's internal guidelines and its public-facing policies. Using its unique position to **ask questions** of the company, the Oversight Board established that the internal reviewer guidance on sexual solicitation differed from Meta's written policy in a way that resulted in the incorrect removal of permitted content and had disproportionate negative impacts on women and gender diverse users.

This decision also highlights the notable limitations of the Oversight Board in effecting policy change. Numerous public comments and the Oversight Board itself raised serious concerns about Meta's *Adult Nudity and Sexual Activity* and *Sexual Solicitation* policies. These concerns included the confusing combination of multiple prohibitions and exceptions in the relevant policies, and the distinction between 'male' and 'female' nipples which disproportionately limits women's expression and presumptively sexualises 'female' nipples. The Oversight Board specifically found that Meta's policies create 'greater

---

[10] Bartolo, Louisa. 2021. '"Eyes Wide Open to the Context of Content": Reimagining the Hate Speech Policies of Social Media Platforms through a Substantive Equality Lens.' Renewal: A Journal of Social Democracy 29(2): 39-51.

[11] Suzor, Nicolas, Sarah Myers West, Andrew Quodling and Jillian York. 2019. 'What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation', International Journal of Communication 13: 1526-1543.

[12] Oversight Board. 2022. 'Gender identity and nudity. https://www.oversightboard.com/decision/BUN-IH313ZHJ/. 'Sexual solicitation.' Accessed February 23, 2023. https://transparency.fb.com/en-gb/policies/community-standards/sexual-solicitation/.

[13] Meta. 'Sexual solicitation.' Accessed February 23, 2023. https://transparency.fb.com/en-gb/policies/community-standards/sexual-solicitation/.

barriers to expression for women, trans and gender non-binary people.' However, while the Oversight Board can make recommendations to Meta, it **does not have authority to make binding decisions about policy changes.** As the Oversight Board is a decision-reviewing body, this may be appropriate from a democratic, separation of powers perspective, but leaves the issue of discriminatory policies unaddressed. **Further research** is therefore needed into how platforms should develop their policies to promote gender equality, including whether a **different model of SMC** may be appropriate for this purpose. This is a particularly difficult challenge in relation to content, including female nudity, that is acceptable in some cultures, contexts and locations, but not others.

Conclusion

This case study has revealed three main benefits of the Oversight Board in addressing gender inequality: (1) exposure of gender inequality in content moderation; (2) greater transparency around complex decisions; and (3) greater public involvement in decision-making.

These benefits have been facilitated by the following features of the Oversight Board, which should be adopted in the establishment of future SMCs: opportunity for public comment; sufficient independence from Meta; ability to ask Meta questions about its policies and practices; publicly available reasoning; access to independent research; and an external set of rules to guide decision-making. The Oversight Board has recently named gender as one of its strategic priorities,[14] and any relevant future decisions should be analysed for further insights.

Although this study has focused on Meta's Oversight Board, the main benefits and features identified are not platform-specific. Provided they have adequate funds, any platform that performs content moderation could implement a SMC with similar features and would likely see similar benefits. These findings could also be applied to other SMC models, including multi-platform and multi-stakeholder SMCs.

Similarly, although this paper has focused on the issue of gender inequality, the benefits and features of the Oversight Board are not gender-specific and may be applicable to other types of structural inequality. Analyses of decisions relating to other types of inequality should be undertaken to investigate the potential of SMCs in these contexts.

This case study also highlighted a significant limitation of the Oversight Board: as a decision-reviewing body, it cannot directly address gender inequality in content moderation policies and enforcement practices. Further research is therefore needed into whether the Oversight Board, or a different model of SMC, should be empowered to make binding decisions about platform policies and practices.

---

[14] Oversight Board. 2022. 'Oversight Board announces seven strategic priorities.' https://www.oversightboard.com/news/543066014298093-oversight-board-announces-seven-strategic-priorities/.

Executive-appointed social media councils: A case study from India

Prateek Waghre
INTERNET FREEDOM FOUNDATION, NEW DELHI, INDIA

Abstract

This short paper presents a case study of the executive-appointed social media councils in India. First, the paper presents the context surrounding the subordinate legislation that enabled the creation of these councils. It proceeds to critique these councils based on legitimacy, lack of specificity and capacity. The paper hypothesises that in their current form, the councils may counterbalance platform power but, in the process, concentrate power in the hands of 'old school speech regulation' bodies.

Background

This section provides the context in which the rules that enabled the formation of the executive-appointed councils were notified. It highlights the tensions between certain social media platforms and the executive branch of the government, as well as the efforts of the latter to exert control over the internet.

In February 2021, two ministries of the Government of India (GoI), the Ministry of Electronics and Information Technology (MeitY) and the Ministry of Information and Broadcasting (MIB), held a joint press conference announcing the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (IT Rules, 2021).[1] The notified version of the IT Rules, 2021 would go into effect on 25th May, 2021, and were significantly different from the draft Information Technology [Intermediaries Guidelines (Amendment) Rules] 2018 (Intermediary Guidelines, 2018) that were open for public feedback in December, 2018. The Intermediary Guidelines were made available for consultation after press reports of close-door meetings and a draft version being published by civil society organisations.[2]

The IT Rules 2021 divided administration between MeitY and MIB. MeitY would administer Part II of the rules, pertaining to 'Due Diligence by Intermediaries and Grievance Redressal Mechanism(s)'. Part II of the rules included obligations that intermediaries would have to fulfil; the mechanism for a grievance redressal process which included requirements to appoint a Grievance Redressal Officer, timelines for acknowledgement and disposal of grievances. Part II also defined a new category of intermediaries, called significant social media intermediaries if the number of registered users in India

---

[1] IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 English, 25 February 2021, https://mib.gov.in/sites/default/files/IT%28Intermediary%20Guidelines%20and%20Digital%20Media%20Ethics%20Code%29%20Rules%2C%202021%20English.pdf. Press Information Bureau, India, 'Government Notifies Information Technology (Intermediary Guidelnes and Digital Media Ethics Code) Rules 2021', 25 February 2021, https://pib.gov.in/PressReleseDetailm.aspx?PRID=1700749.

[2] Seema Chisti, 'Govt Moves to Access and Trace All "Unlawful" Content Online', The Indian Express (blog), 24 December 2018, https://indianexpress.com/article/india/it-act-amendments-data-privacy-freedom-of-speech-fb-twitter-5506572/. 'India Must resist the lure of the Chinese model of online surveillance and censorship #IntermediaryRules #RightToMeme #SaveOurPrivacy, Internet Freedom Foundation, 24 December 2019, https://internetfreedom.in/india-must-resist-the-lure-of-the-chinese-model-of-surveillance-and-censorship-intermediaryrules-righttomeme-saveourprivacy/.

were higher than 5 million, as defined via gazette notification in February 2021. Significant social media intermediaries were also required to have in-country Chief Compliance Officers, Grievance Redressal Officers, and Nodal Officers. They were also obliged to enable 'traceability' of messages to the first originator of a message in India.

Part III of the IT Rules, 2021 were to be administered by MIB and were applicable to publishers of 'news and current affairs content', and 'online curated content'. The Rules proposed a three-tier grievance redressal mechanism, with the publisher forming the first level. The second level would consist of self regulating bodies, and the third, an oversight mechanism consisting of an Inter-Departmental Committee with powers to issue guidelines, advisories, order and directions to publishers.

Between January and June 2021, MeitY and Twitter were interlocked in a 'jawboning' exercise. In late January, Twitter had complied with and then reversed course on some content takedown orders issued by MeitY related to the hashtag 'ModiPlanningFarmerGenocide' stating that it would not take actions against accounts belonging to 'news media entities, journalists, activists, and politicians' as it would violate their 'fundamental right to free expression under Indian law'.[3] In early February, the Government of India threatened Twitter with penal action for not complying with content takedown orders.[4] An unnamed government source would claim that Twitter took down '90-95%' of the accounts.[5] Various government officials and ministries also created accounts and advertised their presence on Koo, an India-based microblogging service.[6] In May 2021, Twitter flagged a post by a spokesperson of the Bhartiya Janta Party (BJP) as containing 'manipulated media'.[7] MeitY wrote to Twitter expressing its objection.[8] The Delhi Police issued notice to Twitter India's Managing Director, Manish Maheshwari.[9] This was followed by a special cell of the Delhi Policy arriving at Twitter's office in New Delhi seeking information.[10] In June 2021, with the IT Rules, 2021 going into effect, Twitter's India MD was named in First Information Reports in at least 2 instances.[11]

---

[3] Twitter Safety. 'Updates on Our Response to Blocking Orders from the Indian Government'. 10 Febrzary 2021. https://blog.twitter.com/en_in/topics/company/2020/twitters-response-indian-government; Billy Perigo, 'Twitter Blocks Accounts Linked to India Farmers Protests | Time', 1 February 2021, https://time.com/5935003/india-farmers-protests-twitter/.

[4] Yuthika Bhargava. 'Farmers' Protests | Govt Issues Notice to Twitter on "Farmer Genocide" Hashtag. The Hindu. 3 February 2021. Sec. India, https://www.thehindu.com/news/national/farmers-protest-govt-issues-notice-to-twitter-warns-of-penal-action/article33739720.ece.

[5] Yuvrai Malik. 'Twitter Takes down "90-95%" Accounts in Line with Meity Orders: Govt Source | Business Standard News'. 12 February 2021, https://www.business-standard.com/article/current-affairs/twitter-takes-down-90-95-accounts-in-line-with-meity-orders-govt-source-121021200134_1.html.

[6] Niharika Sharma, 'The Indian Government Is Backing a Homegrown Alternative Because Twitter Won't Bend to Its Will', Quartz, 10 February 2021, https://qz.com/india/1970534/piyush-goyal-other-indian-ministers-promote-koo-app-on-twitter/.

[7] Special Correspondent, 'Twitter Flags Sambit Patra's Tweet on Congress "Toolkit" as Manipulated Media', *The Hindu*, 21 May 2021, sec. India, https://www.thehindu.com/news/national/twitter-flags-sambit-patras-tweet-on-congress-toolkit-as-manipulated-media/article34611486.ece.

[8] Yuthika Bhargava, 'Government Asks Twitter to Remove "Manipulated Media" Tag Fom Tweets Related to "Congress Toolkit"', The Hindu, 21 May 2021, sec. India, https://www.thehindu.com/news/national/government-asks-twitter-to-remove-manipulated-media-tag-fom-tweets-related-to-congress-toolkit/article34615696.ece.

[9] Live Law [@LiveLawIndia], 'Delhi Police on 21st May Issued Notice to Twitter MD Requesting His Presence on 22nd May in Connection with the "Congress Tool Kit" Matter @TwitterIndia @DelhiPolice #ToolkitCase @INCIndia Https://T.Co/RAfHGsRZHW', Tweet, Twitter, 24 May 2021, https://twitter.com/LiveLawIndia/status/1396845158805630987.

[10] 'Police at Twitter's Door after BJP Posts Flagged; Opposition Slams "Intimidation"', The Indian Express (blog), 25 May 2021, https://indianexpress.com/article/india/twitter-india-delhi-police-raid-7328607/.

[11] Ismat Ara, 'Late Night FIR Against Twitter, Opposition Leaders, Journalists for Posts on Ghaziabad Attack', The Wire, 16 June 2021, https://thewire.in/government/late-night-fir-against-twitter-opposition-leaders-journalists-for-posts-on-ghaziabad-attack; Indu Bhan,'Delhi Police Receives Complaint against Swara Bhaskar, Twitter India MD in Ghaziabad Assault Case', *The Economic Times*, 17 June 2021, https://economictimes.indiatimes.com/news/india/delhi-police-receives-complaint-against-swara-bhaskar-twitter-india-md-in-ghaziabad-assault-case/articleshow/83597364.cms.

The IT Rules 2021 were criticised by civil society organisations on the basis of lack of constitutionality, exceeding the scope of the parent Information Technology Act, 2000, variation from the Intermediary Guidelines, 2018 and absence of subsequent public consultation, among others.[12] As of May, 2022, there were at least 17 challenges in various High Courts across India.[13] While the Supreme Court of India stayed proceedings in these cases, any interim orders were to remain in effect.[14] Certain clauses of Part III of the rules have been stayed by the High Courts of Kerala, Bombay and Madras. However, information revealed in response to Right to Information requests revealed that over 2000 news publishers had furnished details to MIB, even though the 3-tier mechanism had been stayed by the Bombay High Court. Facebook and Whatsapp have also challenged the traceability requirements before the Delhi High Court.[15]

Executive-appointed social media councils

In spite of the various challenges to the IT Rules, 2022, on 3rd June 2022, MeitY proposed, withdrew a set of amendments to the IT Rules 2021.[16] On 6th June 2022, it once again published the proposed amendments for public feedback prefaced by a Press Note which stated that 'early stage or growth stage Indian companies or Startups' would not be impacted, without specifying how.[17] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022 (IT Amendment Rules, 2022) were officially notified in October 2022.[18] The amendments proposed the creation of one or multiple Grievance Appellate Committees (GACs) where users could appeal against decisions taken by the Grievance Officer of any intermediary, with decision of the GACs being binding. A GAC will consist of one chairperson and 2 full time members, and one of these members will be a member ex-officio. A GAC shall 'endeavour to resolve the appeal finally within thirty calendar days'. In June 2022, the Minister of State of Electronics and Information Technology (MoS-EIT) stated that despite appointing grievance officers intermediaries were not providing 'real redressal', which needed

---

[12] Archana Sivasubramanian and Manish, 'Unpacking the IT Rules, 2021', CPR (blog), 21 December 2021, https://cprindia.org/unpacking-the-it-rules-2021/; 'ANALYSIS OF THE INFORMATION TECHNOLOGY (INTERMEDIARY GUIDELINES AND DIGITAL MEDIA ETHICS CODE) RULES, 2021', SFLC.in, 27 February 2021, https://sflc.in/analysis-information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021/; Neeti Biyani and Amrita Choudhury, 'Internet Impact Brief: 2021 Indian Intermediary Guidelines and the Internet Experience in India', *Internet Society* (blog), accessed 12 February 2023, https://www.internetsociety.org/resources/2021/internet-impact-brief-2021-indian-intermediary-guidelines-and-the-internet-experience-in-india/; 'How the Intermediaries Rules Are Anti-Democratic and Unconstitutional.', Internet Freedom Foundation, 27 February 2021, https://internetfreedom.in/intermediaries-rules-2021/.

[13] 'Table Summarizing Challenges to IT Rules, 2021', Google Docs, accessed 12 February 2023, https://docs.google.com/document/d/1kmq-AIRO1XpPaThvesI5xQq2nVkZv6UdmaKFAJ8AMTk/edit?usp=embed_facebook.

[14] 'Supreme Court Stays Proceedings before High Courts Challenging IT Rules, 2021, Interim Orders to Continue', Internet Freedom Foundation, 9 May 2022, https://internetfreedom.in/supreme-court-stays-proceedings-before-high-courts-challenging-it-rules-2021-interim-orders-to-continue/.

[15] PTI, 'WhatsApp Challenges New IT Rules in Delhi HC, Terms It "Unconstitutional"', 26 May 2021, https://theprint.in/india/whatsapp-challenges-new-it-rules-in-delhi-hc-terms-it-unconstitutional/666023/.

[16] 'MeitY Abruptly Withdraws Proposed Amendments to Technology and Social Media Rules - ET Government', ETGovernment.com, 3 June 2022, https://government.economictimes.indiatimes.com/news/governance/meity-abruptly-withdraws-proposed-amendments-to-technology-and-social-media-rules/91974760; 'MeitY Publishes and Then Withdraws a Proposal to Amend IT Rules, 2021', Internet Freedom Foundation, 3 June 2022, https://internetfreedom.in/meity-publishes-and-then-withdraws-a-proposal-to-amend-it-rules-2021/.

[17] Ministry of Electronics and Information Technology, 'Press Note Dated 6 June 22 and Proposed Draft Amendment to IT Rules 2021', 6 June 2022, https://www.meity.gov.in/writereaddata/files/Press%20Note%20dated%206%20June%2022%20and%20Proposed%20draft%20amendment%20to%20IT%20Rules%202021.pdf.

[18] Ministry of Electronics and Information Technology, 'Notification Dated, the 28th October, 2022 G.S.R. 794(E): The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022 | Ministry of Electronics and Information Technology, Government of India', 28 October 2022, https://www.meity.gov.in/content/notification-dated-28th-october-2022-gsr-794e-information-technology-intermediary-guidelines.

to be addressed.[19] Repeating this position in October 2022, the MoS-EIT added that the IT Amendment Rules, 2022 were meant to make the internet safer and that the government did not want to be ombudsmen for the internet and was doing so reluctantly.[20]

The Asia Internet Coalition proposed self-regulatory mechanisms instead of the GAC-model.[21] According to media reports, U.S.-India Business Council (USIBC), part of the U.S. Chamber of Commerce, and U.S.-India Strategic Partnership Forum (USISPF) internally discussed concerns such as independence of the GACs, lack of checks and balances, absence of civil society representation.[22] Civil society organisations and experts raised concerns about executive influence on GACs and their binding decisions about content, their suspect legality on account of being outside the IT Act, 2000, with some even calling for their withdrawal.[23] In January 2022, MeitY invited applications for full time membership of the GACs with a deadline of 12th January, 2022.[24] On 27th January, 2022 it notified the formation of 3 3-member GACs with each being chaired by a member of the Ministries of Home Affairs, Information and Broadcasting, Electronics and Information Technology, respectively.[25] An accompanying press release stated that committees will be functional from 1st March, 2023 after which users would be able to "appeal against (the) decision of the grievance officer of the social media intermediaries and other online intermediaries."[26]

---

[19] PTI, 'Social Media Platforms Not Adequately Redressing Grievances: Rajeev Chandrasekhar', Business Today, 23 June 2022, https://www.businesstoday.in/latest/policy/story/social-media-platforms-not-adequately-redressing-grievances-rajeev-chandrasekhar-338953-2022-06-23.

[20] Yuthika Bhargava, 'Amended IT Rules Are Meant to Make Web Safer for All: Rajeev Chandrasekhar', The Hindu, 29 October 2022, sec. Interview, https://www.thehindu.com/opinion/interview/govts-intent-is-to-work-with-social-media-intermediaries-not-be-seen-as-adversarial-it-mos/article66070407.ece; Aihik Sur and Deepsekhar Choudhury, 'IT Rules Amendment: Govt Doesn't Want to Be Internet Ombudsman, Says Rajeev Chandrasekhar', Moneycontrol, 31 October 2022, https://www.moneycontrol.com/europe/?url=https://www.moneycontrol.com/news/business/startup/grievance-panel-for-intermediaries-will-be-a-traffic-signal-for-user-appeals-rajeev-chandrasekhar-9417991.html; Aditi Agrawal, '"Doing This Reluctantly", Redressal Mechanism "Broken": IT Minister Explains Need for Changed Rules', Newslaundry, 29 October 2022, https://www.newslaundry.com/2022/10/29/doing-this-reluctantly-redressal-mechanism-broken-it-minister-explains-need-for-changed-rules.

[21] Anushka Jain, 'IT Rules 2021: Submission on Safe Harbour Status, Enforcing Compliance, Other Issues', MediaNama (blog), 2 August 2022, https://www.medianama.com/2022/08/223-asia-internet-coalition-submission-on-it-rules-gac-compliance/.

[22] Aditya Kalra and Munsif Vengattil, 'U.S. Lobby Groups Cast Doubts over Independence of India Content Appeal Panel | Reuters', 20 July 2022, https://www.reuters.com/technology/us-lobby-groups-cast-doubts-over-independence-india-content-appeal-panel-2022-07-20/.

[23] Namrata Maheshwari Chima Raman Jit Singh, 'Civil Society Calls on Indian Government to Withdraw Amendments to IT Rules', Access Now (blog), 12 July 2022, https://www.accessnow.org/press-release/india-it-rules-amendments-joint-submission/; Aarathi Ganesan, 'Experts Flag Free Speech and Self-Censorship in India's Amended IT Rules', MediaNama (blog), 22 June 2022, https://www.medianama.com/2022/06/223-it-rules-amendments-india-free-speech-big-tech/; Tejasi Panjiar, 'A Public Brief on the IT Amendment Rules, 2022 a.k.a "How the Government Is Trying to Moderate Online Speech"', Internet Freedom Foundation, 10 November 2022, https://internetfreedom.in/public-brief-on-the-it-amendment-rules-2022/.

[24] Ministry of Electronics & IT [@GoI_MeitY], '#Hiring! Applications Are Invited for the Appointment of Full-Time Members of GAC(s). To Apply, Send an Email with a Scanned Copy of Your Duly Completed Application Form and CV to Group Coordinator (Cyber Law Division), @GoI_MeitY at Cyberlaw-Legal@meity.Gov.in. #DigitalIndia Https://T.Co/C6hp5cM3iH', Tweet, Twitter, 3 January 2023, https://twitter.com/GoI_MeitY/status/1610143208000081926; Digital India [@_DigitalIndia], '#Hiring! Applications Are Invited for the Appointment of Full-Time Members of GAC(s). To Apply, Send an Email with a Scanned Copy of Your Duly Completed Application Form and CV to Group Coordinator (Cyber Law Division), @GoI_MeitY at Cyberlaw-Legal@meity.Gov.in. #DigitalIndia Https://T.Co/R3RLwtLUx4', Tweet, Twitter, 3 January 2023, https://twitter.com/_DigitalIndia/status/1610143064076742659.

[25] Ministry of Electronics and Information Technology, 'Establishment of Grievance Appellate Committees under Rule 3A of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 | Ministry of Electronics and Information Technology, Government of India', accessed 13 February 2023, https://www.meity.gov.in/content/establishment-grievance-appellate-committees-under-rule-3a-information-technology.

[26] Press Information Bureau, India, 'Three Grievance Appellate Committees (GACs) Notified on the Recently Amended "IT Rules 2021"', 28 January 2023, https://pib.gov.in/pib.nic.in.

Discussion and recommendations

This section critiques the GACs on the basis of legitimacy, lack of specificity and capacity. It considers aspects such as the consultation process, legislative uncertainty, operational transparency, vagueness of remit as well as the ability, in terms of scale and expertise, to adjudicate the volume and likely complex nature of appeals expected through the process.

Legitimacy

The GACs draw their basis from a set of rules that had a significant impact on how people in India interact with services on the internet. Therefore the adherence to an open, transparent and responsive public consultation process should have been a crucial component of their drafting. However, the Intermediary Rules, 2018 were released for public consultation after media reports of closed-door meetings, and a version being released in the public domain. These rules did not explicitly seek the appointment of grievance redressal officers, nor oversight of 'news and current affairs content', and 'online curated content'. The IT Rules, 2021 were announced and went into effect in 3 months without any public consultation in their final form.

In March 2021, the Global Network Initiative (GNI) had written a letter to the then Union Minister for Electronics and Information Technology calling on MeitY to 'consider revising the rules and engage in an open, deliberative process about how to address and mitigate these concerns'.[27] Civil society organisations flagged that the process was in contravention of the Pre-legislative Consultation Policy, 2014, lacked a principled approach and were instead driven by political economy, and even called for their withdrawal.[28] Instead of engagement, the minister stated in an interview that they were based on prior consultations, committee reports and court rulings.[29]

Further, multiple civil society organisations also expressed the position that the IT Rules, 2021 were unconstitutional, and went beyond the scope of the parent act.[30] Since the IT Amendment Rules, 2022 are based on the IT Rules, 2021, this assessment extends to them. These concerns were reemphasised specifically with regard to the GACs when they were proposed and subsequently notified stating that there was legislative uncertainty as an executive-appointed council could decide on matters related to free speech based on grounds that were not stated under Section 69A of the IT Act, 2000 or Article 19(2) of the Constitution of India.[31]

---

[27] 'GNI Analysis: Information Technology Rules Put Rights at Risk in India', Global Network Initiative, accessed 14 February 2023, https://globalnetworkinitiative.org/india-it-rules-2021/.

[28] 'ANALYSIS OF THE INFORMATION TECHNOLOGY (INTERMEDIARY GUIDELINES AND DIGITAL MEDIA ETHICS CODE) RULES, 2021'; Archana Sivasubramanian and Manish, 'Unpacking the IT Rules, 2021'; 'Dear MEITY, Withdraw the New IT Rules!', Internet Freedom Foundation, 23 March 2021, https://internetfreedom.in/withdraw-the-it-rules/.

[29] Aashish Aryan, "'Our Commitment to Privacy Is Unimpeachable. Are They Permitting Free Speech by Not Obeying Constitution?': Ravi Shankar Prasad', The Indian Express (blog), 29 May 2021, https://indianexpress.com/article/business/economy/ravi-shankar-prasad-it-rules-privacy-twitter-whatsapp-7334837/.

[30] 'How the Intermediaries Rules Are Anti-Democratic and Unconstitutional.'; Torsha Sarkar et al., 'On the Legality and Constitutionality of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021', The Centre For Internet & Society, 21 June 2021, https://cis-india.org/internet-governance/legality-constitutionality-il-rules-digital-media-2021.

[31] Ganesan, 'Experts Flag Free Speech and Self-Censorship in India's Amended IT Rules'; Tejasi Panjiar, 'A Public Brief on the IT Amendment Rules, 2022 a.k.a "How the Government Is Trying to Moderate Online Speech"'.

Lack of specificity and capacity

Table 1 collates the approximate number of decisions taken by a specific set of significant social media intermediaries based on user reports and self reported numbers for proactive action or action taken. Between October and December 2022, Sharechat received nearly 8 million user reports.[32] In the same period Koo received nearly 30000 user reports and took more than 80,000 proactive content moderation decisions, Facebook and Instagram took action against over 65 million and nearly 10 million pieces of content respectively, and Snap Inc. reported receiving over 550,000 content and account reports.[33] If appeals against even 0.1% of these actions make it to the GACs, they would have to deal with tens of thousands appeals on a monthly basis. Meta's oversight board, with an operating budget of over 100 million dollars has received over 2 million appeals and picked up only 42 cases.[34]

Approximate number of decisions taken based on user reports and action taken disclosures

Social Media Platform	October 2022	November 2022	December 2022
Facebook	over 27 million	over 18 million	over 21 million
Instagram	over 2.6 million	over 3.1 million	over 3 million
Sharechat	over 3.9 million	over 1.8 million	over 2.1 million
Snap Inc.	over 195,000	over 180,000	over 175,000
Koo	over 29000	over 34000	over 16000

Table 1: Compiled by author based on IT Rules, 2021 compliance disclosures.

The press release announcing the constitution of 3 GACs stated that they would 'endeavour to resolve the appeal finally within thirty calendar days'. It does not explicitly state whether the GACs will adjudicate all the appeals or have discretion over which cases they choose. If so, there are immediate questions about their ability to handle the potential volume of cases. Alternatively, if the GACs are meant to have discretion over which cases to pick, then any guiding criteria for doing so have not been defined, thus also raising the risk of arbitrary or motivated case selection. The presence of ex-officio members, who are each currently chairpersons of their respective GACs raise questions about the independence of the GACs.

While the press release states that '(p)eriodic reviews of GACs and reporting and disclosures of GAC orders' will be a part of the process, there has been no further information about the specifics of what these reviews and disclosures will contain, nor the frequency of any such reports. In addition, no operating budgets have been specified. While the conversation mainly revolves around social media intermediaries, the GACs will also be the point of appeal for decisions taken by grievance redressal officers across all kinds of intermediaries. This vagueness of remit creates potential for the GACs to make content moderation decisions at different layers of the technology stack as well.

There is also no clarity on the basis of allocation of appeals across the 3 GACs and what implications that may have for the grievance filing process. Though the IT Amendment Rules, 2022 allow for the

---

[32] 'ShareChat Transparency Reports', accessed 14 February 2023, https://help.sharechat.com/transparency-report/.

[33] 'Koo: View the Latest Koos » Monthly Compliance Reports', accessed 14 February 2023, https://info.kooapp.com/monthly-compliance-reports/; 'Regulatory and Other Transparency Reports | Transparency Center', accessed 14 February 2023, https://transparency.fb.com/data/regulatory-transparency-reports/; 'India Transparency & Data | Snapchat Transparency', accessed 14 February 2023, https://values.snap.com/en-GB/privacy/transparency/india.

[34] Steven Levy, 'Inside Meta's Oversight Board: 2 Years of Pushing Limits | WIRED', WIRED, 8 November 2022, https://www.wired.com/story/inside-metas-oversight-board-two-years-of-pushing-limits/; 'The Oversight Board | Transparency Center', accessed 14 February 2023, https://transparency.fb.com/en-gb/oversight/.

assistance from person(s) 'having requisite qualification, experience and expertise in the subject matter', the inadequate representation of civil society, academia, professionals with expertise in trust and safety positions, etc. mean that GACs may not be adequately equipped to adjudicate what are likely a high volume of complex appeals.

Implications and recommendations

The proposed binding nature of the decisions combined with the general lack of transparency and specificity in a contested, polarised sphere such as social media leave the mechanism vulnerable to executive and ideological capture. The overall approach of the GACs also reflects an approach to content moderation that is neither suitable nor capable of scaling to meet the many challenges in today's information ecosystem. It relies on highly context-specific decisions taken about individual pieces of content, which may or may not have any precedent-setting value, to attempt to address systemic issues that are caused by broader societal-level problems. Aggregation of individual decisions may not be able to address underlying problems since they are neither repeatable nor broadly applicable, given the complexities involved.[35]

Based on the approach to consultative processes, the overreach of subordinate legislation with the IT Rules, 2021 and subsequent amendments; low levels of operational transparency; vagueness of remit across multiple bodies and kinds of intermediaries; as well as questions surrounding the committees in terms of ability, from the perspectives of scale and capability; strong and sustained adherence to principles like human rights, rule of law and democratic values is unlikely. For these reasons, in the Indian context, an executive-appointed and thereby executive-controlled council is an undesirable intervention.

Due to the expected volumes, context-specific nature of content reports and disputes on social media mean even councils that are multi stakeholder in nature can be expected to find it difficult to address problems at scale. Decision-making for councils is likely to be further complicated by the tendency of adversarial groups to employ techniques like malign creativity to evade detection, accountability and introduce a layer of plausible deniability.[36] For example, a combination of the names of two taxi aggregators in India are often employed as an anti-minority dog whistle.[37] There is a need to better understand the impact of social media-based communication on collective behaviour.[38] Thus even multi-stakeholder councils should be approached conservatively, and not result in diversion of resources for appropriate research into understanding their impact.

---

[35] Evelyn Douek, 'Content Moderation as Systems Thinking', *Harvard Law Review* 136, no. 2 (2022), https://harvardlawreview.org/2022/12/content-moderation-as-systems-thinking/.

[36] Nina Jankowicz et al., 'Malign Creativity: How Gender, Sex, and Lies Are Weaponized Against Women Online | Wilson Center', accessed 17 February 2023, https://www.wilsoncenter.org/publication/malign-creativity-how-gender-sex-and-lies-are-weaponized-against-women-online; Mohua Das, 'How Desi Troll Armies Have Built a Coded Language of Abuse | India News - Times of India', 8 February 2022, https://timesofindia.indiatimes.com/india/how-desi-troll-armies-have-built-a-coded-language-of-abuse/articleshow/89406114.cms; Aishwarya Varma, 'Can Tech and Humans Work Together To Make Social Media Less Communally Charged?', TheQuint, 27 April 2022, https://www.thequint.com/news/webqoof/communal-language-and-moderation-social-media-india.

[37] Prateek Waghre, prateekwaghre@mastodon.social [@prateekwaghre], 'It Amazes Me That the Names of the Two Leading App-Based Cab Aggregators in India Have Been Combined in Word Play That Serves a Bigoted End. Malign Creativity" Indeed'. Tweet, *Twitter*, 18 August 2021, https://mobile.twitter.com/prateekwaghre/status/1428011618105913347.

[38] Joseph B. Bak-Coleman et al., 'Stewardship of Global Collective Behavior', *Proceedings of the National Academy of Sciences* 118, no. 27 (6 July 2021): e2025764118, https://doi.org/10.1073/pnas.2025764118.