

# AUTHORITARIANISM AS A SERVICE: INDIA’S MOVES TO WEAPONIZE PRIVATE SECTOR CONTENT MODERATION WITH THE 2021 INFORMATION TECHNOLOGY RULES

*Michael Karanicolas\**

**ABSTRACT** *A central regulatory challenge related to the spread of Internet access is that the power of mass communication has both been democratized and decentralized into the hands of anyone with a connected device, but also consolidated in a handful of massive tech companies, providing opportunities for unprecedented surveillance and control over the public discourse. This paper argues that the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 pose a severe threat to freedom of expression in India, by providing an avenue for the government to coopt these enormous powers under revised safe harbor provisions that essentially turn private sector content moderation powers into a cudgel to target government critics outside of traditional mechanisms of constitutional scrutiny. While India is not the first country to adopt a “jawboning” strategy against platforms to suit domestic political purposes, this posture goes far beyond what any comparable democratic country has attempted. Ultimately, the paper argues in favor of India’s potential to assume global leadership in fostering robust public accountability in platform governance, but that this requires the country to put human rights and democracy at the center of its reform agenda.*

I. Introduction . . . . .	2	IV. Finding Accountability . . . . .	14
II. Regulating India’s Digital Space . . . . .	4	V. Conclusion . . . . .	16
III. The 2021 Rules. . . . .	10		

---

\* Executive Director, UCLA Institute for Technology, Law & Policy. With sincere thanks to Karthik Rai, Sarthak Wadhwa, and Alessia Zornetta for their invaluable research support, as well as to Bharath Gururagavendran, Jyoti Panday, Ayesha Khan, and Akriti Gaur for their helpful feedback and commentary.

## I. INTRODUCTION

Among the thorniest challenges which governments around the world are grappling with are concerns about how to regulate a digital information ecosystem which has evolved beyond what even the most perceptive and forward-thinking constitutional theorists from previous generations could have ever imagined. The change is not merely that information moves exponentially faster, but that the power to speak to a mass audience has been decentralized, from the purview of a small number of elite and powerful channels into the hands of anyone with a connected device. In contrast to the traditional adage that “freedom of the press is guaranteed only to those who own one”,<sup>1</sup> today a simple, free-to-access account on a major platform grants ordinary people with the potential to address an audience in the millions, and to direct their ideas, or their grievances, to the highest echelons of power.<sup>2</sup>

The flip-side to this incredible expansion in expressive opportunity, is that access to the means of mass communication is consolidated in the hands of a small number of international tech companies.<sup>3</sup> These companies, while global in their scope and influence, have traditionally maintained a notoriously myopic attitude to how they manage online speech, dominated by American First Amendment values and American media narratives.<sup>4</sup> Despite the companies’ massive footprint across the developing world, stakeholders in places like Myanmar or the Philippines have struggled to get Facebook, in particular, to pay attention to its often catastrophic impacts on the political discourse.<sup>5</sup>

In the absence of careful attention by the major platforms to these impacts, the migration of more and more of our lives into the online space has given

---

<sup>1</sup> AJ Liebling, ‘Do you Belong in Journalism?’ *The New Yorker* (14 May 1960) 105.

<sup>2</sup> See *eg*, Katie Rogers, ‘Barack Obama Surprises Internet with Ask Me Anything Session on Reddit’ *The Guardian* (29 August 2012) <<https://www.theguardian.com/technology/us-news-blog/2012/aug/29/barack-obama-ask-me-anything-reddit>> accessed 14 July 2022.

<sup>3</sup> Christopher S Yoo, ‘When Antitrust Met Facebook’ (2012) 19 *Geo Mason L Rev* 1147, 1148-58

<sup>4</sup> Chinmayi Arun, ‘Facebook’s Faces’ 135 *Harvard Law Review Forum*; Kate Klonick, ‘The New Governors: The People, Rules, and Processes Governing Online Speech’, (2018) 131 *Harv L Rev* 1598, 1618–21; Michael Karanickolas, ‘Squaring the Circle Between Freedom of Expression and Platform Law’, (2020) 20 *Pitt J Tech L & Pol’y* 175, 183-4.

<sup>5</sup> Eric Johnson, ‘Memo from a ‘Facebook Nation’ to Mark Zuckerberg: You Moved Fast and Broke our Country’ (*Vox*, 11 December 2018) <<https://www.vox.com/2018/11/26/18111859/maria-ressa-rappler-facebook-mark-zuckerberg-philippines-kara-swisher-recode-decode-podcast>>. See also <<https://progressivevoicemyanmar.org/wp-content/uploads/2018/04/Burmese-NGOs-to-Facebook-April-2018.pdf>> accessed 14 July 2022.

rise to unprecedented opportunities for surveillance, manipulation, and ultimately control.<sup>6</sup> A century ago, the vast majority of expression was informal and transient. Even the most despotic king could never hope to control what their subjects actually said about them in the marketplace. Today, nearly every meaningful interaction includes at least some online component, which means that governments which exercise authority over the infrastructure and services for online communication can potentially control virtually the entirety of the national discourse, from what we say about our elected representatives, to how we order our evening meals.<sup>7</sup> Much has been written about the danger posed by massive tech companies, and the harmful impacts of their data-centric business models.<sup>8</sup> However, without discounting these concerns about the surveillance economy, the threat posed by governments wielding these same tools to control their populations is objectively more dangerous than any commercial actor, due to the State's vastly greater coercive power. A new hybrid model of despotism has emerged, where leaders coopt private sector communication networks, and sometimes even commandeer the social media companies themselves, in order to maintain their grip on power.<sup>9</sup>

This paper considers India's latest moves to regulate the digital space in the context of this broader trend towards blurring the line between private sector enforcement and State restrictions on speech to assess the likely impact of the *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* [the 2021 Rules] on freedom of expression in India. While certain aspects of the 2021 Rules address real deficits in accountability related to the consolidation of private sector power, the changes as a whole are structured in a manner which will coopt the platforms' power in support of tightening the government's grip on India's political discourse, to

---

<sup>6</sup> Michael Karanicolas, *Travel Guide to the Digital World - Surveillance and International Standards* (Global Partners Digital 2014, 1, 7-9).

<sup>7</sup> David Moser, *Press Freedom in China under Xi Jinping* (Tina Burrett & Jeff Kingston (eds), Routledge 2019) 68.

<sup>8</sup> See Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Public Affairs, 2019).

<sup>9</sup> For a particularly extreme example of the hand-in-glove relationship between online platforms and State repression, see: Xin Xu, 'Permission Denied: Sina Weibo Behind China's Great Firewall' (2016) 13 *Michigan Journal of Public Affairs* 47. However, while most governments are not as heavy handed as China, the cooption of private sector tools as a mechanism for violating the right to freedom of expression is not an isolated phenomenon, including among democracies. See, eg, Michael Karanicolas, 'Subverting Democracy to Save Democracy: Canada's Extra-Constitutional Approaches to Battling "Fake News"', (2019) 17 *Canadian Journal of Law and Technology* 201, 215-22; 'Israel/Palestine: Facebook Censors Discussion of Rights Issues' (*Human Rights Watch*, 8 Oct 2021) <<https://www.hrw.org/news/2021/10/08/israel/palestine-facebook-censors-discussion-rights-issues>> accessed 14 July 2022.

the detriment of freedom of expression and democracy. Ultimately, the paper argues in favor of India's potential to assume global leadership in fostering robust public accountability in this space, but that this requires the country to put human rights and democracy at the center of its reform agenda, including getting its own house in order to prevent further democratic backsliding.

## II. REGULATING INDIA'S DIGITAL SPACE

The rise of privately-owned internet infrastructure companies as a core player in the global expressive ecosystem has profound implications for traditional understandings of speech regulation.<sup>10</sup> Social media companies, in particular, have an enormous ability to both empower their users, and to silence them. Powerful network effects drive consolidation of the online discourse into the hands of a relatively small number of massive platforms, which retain significant discretion regarding which kinds of speech they will amplify, and who they will or will not do business with.<sup>11</sup> A decision by these major companies to “deplatform” a user, or even to change the way their speech is algorithmically amplified, bears enormous consequences for that person's ability to meaningfully express themselves.<sup>12</sup>

The consolidation of such power in private sector hands, combined with the platforms' comparative lack of transparency or public accountability, is the root of some understandable concern.<sup>13</sup> Where governments exercise similar powers over the public discourse, there are typically a range of checks and balances, and clear international standards for legitimate restrictions and controls impacting public expression.<sup>14</sup> The exercise of this scale of private sector power over the global political discourse creates novel challenges from the perspective of human rights and democracy, resulting in significant scholarly debate over the precise meaning of concepts like freedom of expression and transparency in the context of the new “platform law”.<sup>15</sup>

<sup>10</sup> Jack M Balkin, ‘Free Speech is a Triangle’ (2018) 118 Colum L Rev 2011.

<sup>11</sup> Spencer W Waller, ‘Antitrust and Social Networking’ (2012) 90 NCL Rev 1771.

<sup>12</sup> Daphne Keller, ‘Amplification and its Discontents: Why Regulating the Reach of Online Content is Hard’ (2021) 1 Journal of Free Speech Law 229, 230-3.

<sup>13</sup> Evelyn Douek, ‘The Rise of Content Cartels’ (*Knight First Amendment Institute at Columbia University*, 11 February 2020) <<https://knightcolumbia.org/content/the-rise-of-content-cartels>> accessed 14 July 2022. See also Michael Karanicolas, ‘A FOIA for Facebook: Meaningful Transparency for Online Platforms’ (2021) 66 St. Louis University Law Journal 49 (2021).

<sup>14</sup> See, eg, UN Human Rights Comm, ‘General Comment 34, at art 19, Freedoms of Opinion and Expression’ UN Doc CCPR/C/GC/34 (2011) <<https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>> accessed 14 July 2022.

<sup>15</sup> Michael Karanicolas, ‘Squaring the Circle Between Freedom of Expression and Platform Law’ (2020) 20 Pitt J Tech L & Pol’y 175.

However, the ambiguity around who, exactly, is responsible for governing online speech leaves an opening for governments, some of whom have been keen to weaponize the platforms' unprecedented powers as an avenue for controlling the online discourse, and of circumventing constitutional limitations on state censorship.<sup>16</sup> A key example of this has been the rise of specialized referral agencies as a mechanism for governments to combat platform-hosted content which is viewed as being problematic, though not necessarily illegal.<sup>17</sup> These agencies, which are typically connected to law enforcement, monitor social media and file takedown requests against content on the grounds that it purportedly violates the private terms of service of the platform.<sup>18</sup> The fact that the decision to remove the content is being made by the platform, rather than the government, can place the decision outside of traditional judicial scrutiny, despite the fact that the governments are initiating the complaint, and that there is often a blatantly coercive subtext underlying these requests.<sup>19</sup> For example, in April 2021 the Israeli Supreme Court denied a petition against Israel's Cyber Unit due to the inability of the complainants to establish conclusive factual grounds for their argument, including whether the actions taken by the platforms were solely the result of the unit's referral.<sup>20</sup>

The nuances of this dynamic between governments, their citizens, and private sector platforms mean that regulations impacting online intermediaries need to be considered exceptionally carefully vis-à-vis their impact on the freedom of expression of their users.<sup>21</sup>

India's first major move towards regulating online intermediaries came with the 2000 *Information Technology Act*, which applied new rules to any entity which "receives, stores or transmits" messages, or otherwise delivers services in respect of these messages.<sup>22</sup> However, a major deficiency of this

---

<sup>16</sup> Daphne Keller, 'Who Do You Sue? State and Platform Hybrid Power Over Online Speech' (*Aegis Series Paper* No. 1902, 29 January 2019) <<https://www.lawfareblog.com/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech>> accessed 14 July 2022, 2.

<sup>17</sup> Michael Karanicolas, 'Squaring the Circle Between Freedom of Expression and Platform Law' (2020) 20 *Pitt J Tech L & Pol'y* 175, 185-6.

<sup>18</sup> *ibid.*

<sup>19</sup> Michael Karanicolas, 'Subverting Democracy to Save Democracy: Canada's Extra-Constitutional Approaches to Battling "Fake News"' (2019) 17 *Canadian Journal of Law and Technology* 201, 215-22.

<sup>20</sup> *Adalah Legal Center for Arab Minority Rights in Israel v. State Attorney's Office – Cyber Department* HCJ 7846/19.

<sup>21</sup> Jack M Balkin, 'How to Regulate (and Not Regulate) Social Media', (2019) SSRN <<https://ssrn.com/abstract=3484114>> accessed 14 July 2022.

<sup>22</sup> *Information Technology Act 2000*, s 2(w).

early framework was its lack of proper intermediary liability protections,<sup>23</sup> which left the platforms potentially open to direct, strict liability regarding some types of harmful content posted by their users.<sup>24</sup> The problem with such a structure is that the enormous scale at which online platforms operate renders meaningful editorial control, of the type that a newspaper or broadcaster exercises, to be functionally impossible.<sup>25</sup> Any platform which features user-generated content on a massive scale needs some avenue for meaningful safe harbor against the negative behaviours of these users in order to operate, which is why intermediary liability protections are generally understood as foundational to the modern digital information economy, and to a robust democratic discourse.<sup>26</sup> This rationale, which underlay the original passage of Section 230 of Communications Decency Act in the United States,<sup>27</sup> also explains why serious reform proposals by critics of that statute typically aim at modest adjustments to the ambit of protection offered, as opposed to a wholesale revocation of safe harbour.<sup>28</sup>

In 2008, the *Information Technology Act* was amended to provide for broader safe harbour protections for intermediaries over material for which they were a mere conduit, provided they exercised appropriate “due diligence” against problematic content.<sup>29</sup> However, these reforms were followed by the promulgation of a set of Rules which granted the government enhanced ability to control the flow of online information, through expanded powers to order intermediaries to block and decrypt user content.<sup>30</sup> In 2011, these powers were further supplemented by the *Information Technology (Intermediaries guidelines) Rules*, which mandated the removal of several vaguely defined categories of problematic content within just 36 hours of receiving a complaint.<sup>31</sup> These rules were criticized for incentiviz-

<sup>23</sup> In particular, s 79 of the law provided conditional immunity only for offences under the Information Technology Act, as opposed to blanket immunity for speech related offences by the platforms’ users. See also *Aneeta Hada v Godfather Travels and Tour (P) Ltd* (2012) 5 SCC 661 [64].

<sup>24</sup> *Avnish Bajaj v State* 2008 SCC OnLine Del 688.

<sup>25</sup> See, eg, *persistent challenges applying fair use doctrine to alleged cases of copyright infringement*: Leron Solomon, ‘Fair Users or Content Abusers? The Automatic Flagging of Non-Infringing Videos by Content ID on YouTube’ (2015) 44 Hofstra L Rev 237.

<sup>26</sup> Chinmayi Arun, ‘Gatekeeper Liability and Article 19(1)(a) of the Constitution of India’, (2014)7 NUJS L Rev 73.

<sup>27</sup> 47 USC § 230 (1996).

<sup>28</sup> See, eg, Danielle Keats Citron & Benjamin Wittes, ‘The Internet will not Break: Denying Bad Samaritans Section 230 Immunity’ (2017) 86 Fordham L Rev 401.

<sup>29</sup> Information Technology Act 2000, ss 79(1), 79(2)(c).

<sup>30</sup> Information Technology (Procedure and Safeguards for Blocking for Access of Information by Public) Rules 2009; Information Technology (Procedure and Safeguards for Interception, Monitoring and Decryption of Information) Rules 2009.

<sup>31</sup> Information Technology (Intermediaries Guidelines) Rules, 2011, rr 3(2)(b), 3(4).

ing aggressive moderation by intermediaries, including of material which would normally be protected speech under India's Constitution.<sup>32</sup> By creating a baseline level of moderation enforcement, with penalties for failing to takedown enough material, but no concomitant incentive against moderating too aggressively, the natural result of this structure is that intermediaries will remove anything which is subject to a colorable complaint, regardless of whether or not the content at issue is actually illegal.<sup>33</sup> It is worth noting that short time frames for response are particularly problematic in this context, since they preclude a meaningful standard of review in light of the volume of content that intermediaries carry.

These problems were addressed in the landmark judgment of *Shreya Singhal v. Union of India*, in which the Court recognized the practical limits of intermediaries' ability to examine the totality of user-generated content that they host, reading down the "actual knowledge" requirements for removing content to be limited to cases where a government or court order had been delivered, and emphasizing that the latter instances must be justified within the confines of Article 19(2) of the Constitution.<sup>34</sup> However, s. 69A of the Information Technology Act, which grants the government power to block any websites, was ultimately upheld, with the Court finding sufficient safeguards around the exercise of this power.

Following this judgment, civil society voices called for a framework to formalize protection for intermediary liability, in recognition of the values which the Court spelled out in *Shreya Singhal*.<sup>35</sup> Such protections would be in line with the robust intermediary liability protections which exist in most healthy democracies.<sup>36</sup> However, in the years following this deci-

<sup>32</sup> Ujwala Uppaluri, 'Constitutional Analysis of the Information Technology (Intermediaries' Guidelines) Rules, 2011', (*Centre for Internet and Society*, 16 July 2012) <<https://cis-india.org/internet-governance/constitutional-analysis-of-intermediaries-guidelines-rules>> accessed 14 July.

<sup>33</sup> Wendy Seltzer, 'Free Speech Unmoored in Copyright's Safe Harbor: Chilling Effects of the DMCA on the First Amendment' (2010) 24 Harv J L & Tech 171.

<sup>34</sup> *Shreya Singhal v Union of India* (2013) 12 SCC 73 [116].

<sup>35</sup> Jyoti Panday, 'The Supreme Court Judgment in *Shreya Singhal* and What it does for Intermediary Liability in India?' (*Centre for Internet and Society*, 11 April 2015) <<https://cis-india.org/internet-governance/blog/sc-judgment-in-shreya-singhal-what-it-means-for-intermediary-liability>> accessed 25 December 2020.

<sup>36</sup> See, eg, Communications Decency Act, 47 USC § 230 (1996); Tambiama Madiaga, 'Reform on the EU Liability Regime for Online Intermediaries: Background on the Forthcoming Digital Services Act' (*European Parliamentary Research Service*, May 2020). <[https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS\\_IDA\(2020\)649404\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS_IDA(2020)649404_EN.pdf)> accessed 14 July; *Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, European Commission (December 15, 2020), <<https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>> accessed 14

sion, governments have continued to test the boundaries of constitutional permissibility in their blocking orders.<sup>37</sup> Civil society, for their part, have struggled to obtain reliable information regarding the exercise of content blocking orders.<sup>38</sup> Subsequent judicial decisions have also narrowed the protections that *Shreya Singhal* purported to apply to intermediaries.<sup>39</sup> Notably, the Supreme Court in both *Sabu George*<sup>40</sup> and *In Re: Prajwala*<sup>41</sup> expressed support for automated and proactive filtering measures to remove prescribed categories of sexual content, despite concerns about collateral damage from these automated measures on legitimate speech and, in particular, regarding their potential discriminatory impacts on marginalized groups.<sup>42</sup>

While proactive filtering has become increasingly normalized, particularly to combat child sexual abuse material (CSAM), these uses are still overwhelmingly dependent on human reviewers to manually identify illegal images for the filters to seek out.<sup>43</sup> Moreover, there are still highly publicized examples of these algorithms getting it wrong.<sup>44</sup> This is despite the fact that CSAM is vastly easier to algorithmically identify than, say, hate speech or defamation, since the latter relies on judgments than are far more subtle

---

July; Vivek Krishnamurthy & Jessica Fjeld, *CDA 230 Goes North American? Examining the Impacts of the USMCA's Intermediary Liability Provisions in Canada and the United States*, (SSRN, 2020) <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3645462](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3645462)> accessed 14 July 2022.

<sup>37</sup> Vrinda Bhandari & Anja Kovacs, 'What's Sex Got to Do with it', (SSRN, 2021) <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3769942](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3769942)> accessed 14 July 2022.

<sup>38</sup> See Jyoti Panday, 'DeitY says 143 URLs have been Blocked in 2015; Procedure for Blocking Content Remains Opaque and in Urgent Need of Transparency Measures', (*Centre for Internet & Society*, 29 April 2015) <<https://cis-india.org/internet-governance/blog/deitysays-143-urls-blocked-in-2015>> accessed 14 July 2022.

<sup>39</sup> See *MySpace Inc v Super Cassettes Industries Ltd* 2016 SCC OnLine Del 6382 [65]; *The Registrar (Judicial), Madurai bench of Madras High Court v Secy to Govt, Union Ministry of Communications, Govt of India* 2017 SCC OnLine Mad 25298.

<sup>40</sup> *Sabu Matthew George v Union of India* (2015) 11 SCC 545.

<sup>41</sup> *Videos of Sexual Violence and Recommendations, In re* (2018) 15 SCC 551.

<sup>42</sup> Rishab Bailey and Vrinda Bhandari, 'Rethinking Legal-Institutional Approaches to Sexist Hate Speech in India', (2021) *IT For Change* <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3792184](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3792184)> accessed 14 July 2022, 4-5.

<sup>43</sup> While automated systems have become increasingly sophisticated at recognizing copies of a proscribed image, they still require a reference point to carry out this task, generally a database of proscribed images which has been assembled by a human. See eg, Nikola Todorovic & Abhi Chaudhuri, 'Using AI to Help Organizations Detect and Report Child Sexual Abuse Material Online', (*The Keyword*, 3 September 2018) <<https://www.blog.google/around-the-globe/google-europe/using-ai-help-organizations-detect-and-report-child-sexual-abuse-material-online>> accessed 14 July 2022; Facebook Security, 'Want to know how Facebook uses PhotoDNA? Read a Recent Blog Post by the Head of our Safety Team', (*Facebook*, 10 August 2011) <<https://www.facebook.com/security/posts/want-to-know-how-facebook-uses-photodna-read-a-recent-blog-post-by-the-head-of-0/234737053237453/>> accessed 14 July 2022.

<sup>44</sup> Zoe Kleinman, 'Fury over Facebook 'Napalm Girl' Censorship' (*BBC News*, 9 September 2016) <<https://www.bbc.com/news/technology-37318031>> accessed 14 July 2022.



and contextual. International human rights standards generally hold that automated content filtering is analogous to prior censorship, and is generally an unjustifiable interference with freedom of expression when imposed via a government mandate.<sup>45</sup>

Although it is important not to overstate the specific impact of the *Sabu George* and *In Re: Prajwala* decisions, they have taken place against the backdrop of a broader and widely noted deterioration of India's freedom of expression climate, particularly with regards to online speech.<sup>46</sup> This includes the country's ongoing status as the global leader in Internet shut-downs,<sup>47</sup> which are generally recognized as a grave violation of human rights, and which are prima facie unjustifiable under international human rights law.<sup>48</sup> There is also ample evidence that the *Information Technology Act's* takedown rules are being applied to facilitate massively overbroad blocking orders, including against multi-purpose services such as Indian Kanoon, Reddit, and Telegram.<sup>49</sup>

However, even more troubling than these formal moves to clamp down on online speech are indications of informal engagement with online platforms, particularly Facebook, with the aim of influencing the social network's moderation policy towards certain Indian content.<sup>50</sup> In parallel with these moves, the present government of India has demonstrated their own

<sup>45</sup> 'Joint Declaration on Freedom of Expression and the Internet', (OSCE, 1 June 2011) <[www.osce.org/fom/99558?download=true](http://www.osce.org/fom/99558?download=true)> accessed 14 July 2022, 68.

<sup>46</sup> Soutik Biswas, 'Why Journalists in India are Under Attack' (*BBC News*, 4 February 2021) <<https://www.bbc.com/news/world-asia-india-55906345>> accessed 14 July 2022; 'Freedom in the World 2021: India, Freedom House' (*Freedom House*, 2021) <<https://freedomhouse.org/country/india/freedom-world/2021>> accessed 14 July 2022; 'World Report 2021: India' (*Human Rights Watch*, 2021) <<https://www.hrw.org/world-report/2021/country-chapters/india>> accessed 14 July 2022.

<sup>47</sup> 'Shattered Dreams and Lost Opportunities: A Year in the Fight to #KeepItOn' (*Access Now*, March 2021) <[https://www.accessnow.org/cms/assets/uploads/2021/03/KeepItOn-report-on-the-2020-data\\_Mar-2021\\_3.pdf](https://www.accessnow.org/cms/assets/uploads/2021/03/KeepItOn-report-on-the-2020-data_Mar-2021_3.pdf)> accessed 14 July 2022.

<sup>48</sup> *UNGA Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, David Kaye, United Nations General Assembly (*Human Rights Council*, 23 April 2020) <<https://undocs.org/A/HRC/44/49>> accessed 14 July 2022.

<sup>49</sup> Torsha Sakar & Gurshabad Grover, 'How India is using its Information Technology Act to Arbitrarily Take Down Online Content', (*Scroll*, 22 June 2022) <<https://scroll.in/article/953146/how-india-is-using-its-information-technology-act-to-arbitrarily-take-down-online-content>> accessed 14 July 2022.

<sup>50</sup> Newley Purnell and Jeff Horwitz, 'Facebook's Hate-Speech Rules Collide with Indian Politics', *Wall Street Journal* (August 14, 2020), <<https://www.wsj.com/articles/facebook-hate-speech-india-politics-muslim-hindu-modi-zuckerberg-11597423346>>; Newley Purnell and Jeff Horwitz, 'Facebook Services are Used to Spread Religious Hatred in India, Internal Documents Show', *Wall Street Journal* (October 23, 2021), <<https://www.wsj.com/articles/facebook-services-are-used-to-spread-religious-hatred-in-india-internal-documents-show-11635016354>>; Kari Paul, Facebook Blocked Hashtag Calling for Narendra

willingness to weaponize the platforms' terms of service as a cudgel against their opponents, through the constitution of "cyber warriors" tasked with flagging content for removal, including a particular focus on government critics.<sup>51</sup> In addition to concerns around the extra-constitutional nature of these enforcement actions, which strip users of the normal procedural protections that are meant to apply to cases where governments suppress speech,<sup>52</sup> this behaviour is also part of a growing trend of political actors employing online "troll armies" to harass opposition, and try and drown out, or otherwise bully them offline.<sup>53</sup> Across much of the world, a similar toolkit has been associated with anti-democratic movements and a general backsliding across emerging democracies.<sup>54</sup>

### III. THE 2021 RULES

In the context of a closing civic space online, the promulgation of the 2021 Rules raises a number of concerns, insofar as they present the clearest signal yet that the present government of India seeks to tighten its control over the online political discourse. The 2021 Rules, which supplant the 2011 Rules, were announced in a press release which declared that they would "empower ordinary users of social media, embodying a mechanism for redressal and timely resolution of their grievance" with a framework that is "progressive, liberal and contemporaneous".<sup>55</sup> The government in particular justified the changes by pointing to proliferating threats online, such as abusive language,

---

Modi to Resign over Pandemic, *The Guardian* (April 28, 2021), <<https://www.theguardian.com/technology/2021/apr/28/facebook-blocked-resignmodi-hashtag-india-coronavirus>>.

<sup>51</sup> Rina Chandran, "Sinister form of Censorship: India's 'Cyber Volunteer' Programme Pits Citizens against Each Other", (*Scroll*, 2 December 2021), <<https://scroll.in/article/1011856/sinister-form-of-censorship-indias-cyber-volunteer-programme-pits-citizens-against-each-other>> accessed 14 July 2022; Rishab Bailey, Vrinda Bhandari and Faiza Rahman, 'Examining the Online Anonymity Debate: How far Should the Law go in Mandating User Identification?', (Data Governance Network, 2021) <<https://datagovernance.org/report/examining-the-online-anonymity-debate-how-far-should-the-law-go-in-mandating-user-identification>> accessed 14 July 2022.

<sup>52</sup> See, eg, UN Human Rights Comm, *Freedoms of Opinion and Expression*, U.N. Doc. CCPR/C/GC/34 (2011), at arts 19, General Comment 34 <<https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>> accessed 15 July 2022.

<sup>53</sup> See, eg, Jonathan Corpus Ong & Jason Vincent A. Cabañes, 'Architects of Networked Disinformation: Behind the Scenes of Troll Accounts and Fake News Production in the Philippines', (2018) UMass Amherst Communication Department Faculty Publication Series; Artur Pericles Lima Monteiro, 'Brazilians are Desperately Fighting against Bolsonaro's Digital Tactics' (*Rest of World*, 2 December 2021) <<https://restofworld.org/2021/brazil-bolsonaro-social-media/>> accessed 15 July 2022.

<sup>54</sup> *ibid.*

<sup>55</sup> 'Government notifies Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021', (*Press Information Bureau*, 2021) <<https://pib.gov.in/PressReleaseDetailm.aspx?PRID=1700749>> accessed 15 July 2022.

defamatory and obscene content, “blatant disrespect to religious sentiments”, revenge porn, misinformation, and the “misuse of social media for settling corporate rivalries in blatantly unethical manner”.<sup>56</sup>

However, the new Rules have raised a number of concerns among human rights groups who have been particularly alarmed by measures aimed at undermining the privacy of online communications.<sup>57</sup> Though some within the tech sector have defended these provisions on nationalistic grounds, claiming that they would ensure that the rules around privacy are drawn by Indian courts rather than American ones, there is little prospect that the resulting framework will provide for more privacy protections than the status quo.<sup>58</sup>

Rule 4(4) also seeks to further entrench the use of automated filtering tools by significant social media intermediaries, including for any content which has previously been subject to a removal order.<sup>59</sup> As noted earlier in this paper, the use of filtering technologies to target heavily contextual categories of content, such as hate speech, raises particular concerns due to persistent challenges in accuracy and reliability. Questions have also been raised regarding whether the 2021 Rules are ultra vires their parent legislation, insofar as they may constitute a regulatory framework which exceeds the

---

<sup>56</sup> *ibid.* It is worth noting that, while some of these are real and legitimate concerns, regulating of speech to prevent “disrespect” of religions is generally not considered justified under international human rights law. *See, eg*, UN Human Rights Comm, ‘Freedoms of Opinion and Expression’, UN Doc. CCPR/C/GC/34 (2011), General Comment 34, at art 51 <<https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>> accessed 15 July 2022.

<sup>57</sup> Although this paper focuses exclusively on concerns related to online content and platform governance, this focus is not intended to gloss over other problems with the 2021 Rules. In particular, Rule 4(2), which requires “Significant Social Media Intermediaries” to maintain a means of identifying the “first originator” of information they host, poses a severe threat to online privacy, as well as to freedom of expression. The lack of a timeframe connected to this requirement is particularly concerning, as the volume of information storage that this would necessitate runs counter to fundamental principles of data minimization, as spelled out by Justice B.N. Srikrishna and the Committee of Experts in their White Paper on Data Protection in India. For a full discussion of concerns with the 2021 Rules, *see eg*, Torsha Sarkar, Gurshabad Grover, Raghav Ahojja, Pallavi Bedi and Divyank Katira, ‘On the Legality and Constitutionality of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021’, (*Medianama*, 22 June 2021), <<https://www.medianama.com/2021/06/223-legality-constitutionality-of-it-rules/>> accessed 15 July 2022.

<sup>58</sup> ‘WhatsApp-India case: ‘They have Clear Double Standards’, TV Mohandas Pai Says’, *The Economic Times* (26 May 2021) <<https://economictimes.indiatimes.com/tech/technology/whatsapp-india-case-they-have-clear-double-standards-tv-mohandas-pai-says/article-show/82969539.cms?from=mdr>> accessed 15 July 2022.

<sup>59</sup> This approach is problematic since, for a number of categories of prohibited content, such as intellectual property violations, their legality can depend on the surrounding context.

scope of the original Information Technology Act.<sup>60</sup> However, from the perspective of online content and platform governance, the most relevant aspect of the new Rules concerns the amended requirements for intermediaries to establish “due diligence”, and therefore qualify for safe harbor protections against liability for their users’ speech.

Of particular note is s. 3(1)(b), which appears to mandate that intermediaries revise their terms of service to prohibit certain categories of speech. Some of these align with traditional categories of legally prescribed content, such as defamation, CSAM, and content which violates intellectual property rights. However, there are also a number of categories of prohibited content which are perfectly legal under Indian law, and whose prohibition would never pass muster under international human rights standards or, for that matter, under the laws of virtually any state with meaningful constitutional protections for freedom of expression.<sup>61</sup> For example, s. 3(1)(b)(ii) mandates that intermediaries prohibit content which is “insulting or harassing on the basis of gender” or which is “racially or ethnically objectionable”. Both of these categories of content go vastly beyond the acceptable contours of hate speech law according to international human rights standards, which typically require both that the statement was intended to generate hatred and that the statement was likely to generate actual harm.<sup>62</sup> A number of other categories in the code are hopelessly vague, including information that “belongs to another person” (s. 3(1)(b)(i)), “is harmful to [a] child” (s. 3(1)(b)(iii)), or which “knowingly and intentionally communicates information which is patently false or misleading in nature but may reasonably be perceived as a fact” (s. 3(1)(b)(vi)). It is worth noting that, according to international human rights standards, the presence of overly vague restrictions on speech is itself a violation of the right to freedom of expression, regardless of whether this discretion is actually exercised abusively.<sup>63</sup>

State pressure against intermediaries to improve their site-specific rules against particular types of speech is a relatively common, albeit problematic,

---

<sup>60</sup> See eg, *Agricultural Market Committee v Shalimar Chemical Works Ltd* (1997) 5 SCC 516; *State of Karnataka v H. Ganesh Kamath* (1983) 2 SCC 402.

<sup>61</sup> In the Indian context specifically, it is relevant to note that the Lok Sabha Committee on Subordinate Legislation, in their 31<sup>st</sup> Report on Rules under the IT Act, recommended in 2013 that ambiguities in the Act’s definitions needed to be addressed: <[http://164.100.47.193/lss-committee/Subordinate%20Legislation/15\\_Subordinate%20Legislation\\_31.pdf](http://164.100.47.193/lss-committee/Subordinate%20Legislation/15_Subordinate%20Legislation_31.pdf)> accessed 15 July 2022.

<sup>62</sup> See eg, *Jersild v Denmark*, No 15890/89, [1994] ECHR 33, 19 EHRR 1; *Erbakan v Turkey*, App No 59405/00 (ECHR, 6 July 2006); *Prosecutor v Nahimana*, Case No. ICTR-99-52-T, Judgement and Sentence (December 3, 2003).

<sup>63</sup> UN Human Rights Comm (n 52) at arts 24, 27

feature of global content moderation debates.<sup>64</sup> But while governments frequently approach the line of demanding that platforms undertake specific, extra-legal enforcement actions, India's move to essentially require that site-specific terms of service meet a particular government-imposed standard in targeting various forms of legal content goes significantly beyond the measures taken in any other constitutional democracy. While, in theory, platforms are free to set terms of service that do not align with these provisions, the importance of intermediary liability protections to their operations means that this is not a realistically available option. The government's power to push for moderation structures that suit partisan objectives is even further bolstered by a new mandate in the 2021 Rules which requires platforms to maintain staff that are resident in India, and could therefore be subject to personal consequences, in addition to any commercial pressures the government might bring to bear on the company.<sup>65</sup>

A system which allows the government to set the rules, to trigger enforcement actions, and to exercise pressure over how complaints are resolved is, for all practical purposes, indistinguishable from a direct government restriction on speech. Any purported firewall in the 2021 Rules between government enforced restrictions on speech and the platforms' own "independent" enforcement efforts is no more than a fig leaf to shield a massive expansion in government control over the contours of acceptable speech from meaningful judicial or constitutional scrutiny. Though the success or failure of this gambit remains to be seen,<sup>66</sup> such an overt attempt to weaponize the incredible power that platforms wield over online expression is deeply troubling for the precedent that it sets in India, and among other emerging democracies. For all the legitimate concerns around the consolidation of private sector power in governing the online discourse, an expansion of unaccountable government power into this space is far more dangerous, since it threatens to permanently close off the political discourse from any expression which the ruling party may find inconvenient.

---

<sup>64</sup> Michael Karanickolas, 'Subverting Democracy to Save Democracy: Canada's Extra-Constitutional Approaches to Battling "Fake News"' (2019) 17 Canadian Journal of Law and Technology 201.

<sup>65</sup> See, in particular, the requirements for maintaining employees resident in India, Vittoria Elliott, 'New Laws Requiring Social Media Platforms to Hire Local Staff Could Endanger Employees' (*Rest of World*, May 2021) <<https://restofworld.org/2021/social-media-laws-twitter-facebook/>> accessed 15 July 2022.

<sup>66</sup> *Agij Promotion of Nineteenonea Media (P) Ltd v Union of India* 2021 SCC OnLine Bom 2938.

#### IV. FINDING ACCOUNTABILITY

None of this is to suggest that government regulation of social media platforms is inherently problematic, or that there is no role for governments in ensuring robust accountability and due process in how platforms manage content. Major social media platforms do a poor job of serving their users, particularly across emerging democracies. Defenders of the 2021 Rules are not wrong to point to the inherent tension between the companies' American character and the global services they provide, or about the double standards in how different rules and requests are treated.<sup>67</sup> Despite having made some strides towards better representation in recent years, content moderation resources at most platforms remain overwhelmingly focused on serving the United States and Western Europe.<sup>68</sup> Moreover, the development and application of the platforms' moderation standards remains opaque and generally unaccountable, with a significant need for better and more clearly demonstrated procedural rigor.<sup>69</sup>

Interestingly, the 2021 Rules also include a number of proposals which aim to address this aspect of the problem, by supporting better transparency and public responsiveness in platform decision-making. For example, section 4(d) requires the publication of monthly reports regarding complaints received and removals that have been executed, as well as information about automated moderation tools that are being deployed. Given the pervasive challenge in obtaining accurate information about content moderation, particularly in terms of how platforms' policies and implementation differ depending on the language or region of the world, this is a welcome development.<sup>70</sup>

---

<sup>67</sup> 'Ravi Shankar Prasad, IT Minister Exclusive Interview, Social Media Giants Vs Govt with Marya Shakil' (CNN-News 18, 2021) <<https://youtu.be/akmIuukWEEg>> accessed 15 July 2022.

<sup>68</sup> Craig Silverman, Ryan Mac & Pranav Dixit, "I Have Blood on My Hands": A Whistleblower Says Facebook Ignored Global Political Manipulation', (*Buzzfeed*, 14 September 2020) <<https://www.buzzfeednews.com/article/craigsilverman/facebook-ignore-political-manipulation-whistleblower-memo>> accessed 15 July 2022; Marwa Fatafta, 'Facebook is Bad at Moderating in English. In Arabic, it's a Disaster', (*Rest of World*, 18 November 2021) <<https://restofworld.org/2021/facebook-is-bad-at-moderating-in-english-in-arabic-its-a-disaster/>> accessed 15 July 2022.

<sup>69</sup> Jonathan Zittrain, 'Three Eras of Digital Governance', (*Völkerrechtsblog*, 27 November 2019) <<https://voelkerrechtsblog.org/three-eras-of-digital-governance/>> accessed 15 July 2022.

<sup>70</sup> See eg, Carlos Cortés and Luisa Fernanda Isaza, 'The New Normal? Disinformation and Content Control on Social Media during COVID-19', (*Centro de Estudios en Libertad de Expresión y Acceso a la Información*, April 2021) <[https://www.palermo.edu/Archivos\\_content/2021/cele/papers/Disinformation-and-Content-Control.pdf](https://www.palermo.edu/Archivos_content/2021/cele/papers/Disinformation-and-Content-Control.pdf)> accessed 15 July 2022.

However, even among the aspects of the Rules which approach these challenges more constructively, there are reasons for concern. Probably the most ambitious and novel aspect of the Rules, from an accountability perspective, is the constitution of a sort of co-regulatory structure for processing complaints related to platforms' moderation decisions. This begins with a requirement for user notification related to adverse moderation actions, and for platforms to designate a human point of contact for receiving complaints related to moderation actions.<sup>71</sup> The latter requirement, in particular, is badly needed, as anyone who has struggled through the labyrinthine processes of attempting to get an account restored can tell you.<sup>72</sup> But even here, there are aspects of the framework which lend themselves to political abuse. For example, a key responsibility of the grievance officer under section 3(2)(b) is to facilitate the removal of pornographic or impersonating content upon receipt of a complaint from the subject of the image. While efforts to combat non-consensual pornography are absolutely a legitimate State interest,<sup>73</sup> the extension of this provision to any "artificially morphed images" opens the door to censoring a huge amount of legitimate political satire.<sup>74</sup> This concern is compounded by the fact that the material is required to be removed within 24 hours, which will effectively prevent any meaningful analysis regarding whether the material may be legitimate speech, such as political satire.

Above these grievance officers, the Rules introduce an ambitious vision for one or more "self-regulatory bodies of publishers", comprising a set of independent experts who perform a range of tasks including hearing appeals against content decisions by the platforms.<sup>75</sup> Such "social media councils" have been a mainstay of expert discussions around how to improve online content governance in recent years, with a number of high profile supporters, particularly David Kaye, the former United Nations Special Rapporteur on Freedom of Expression.<sup>76</sup> However, this aspect of the Rules applies not just

---

<sup>71</sup> The 2021 Rules, ss 3(2)(a), 4(8).

<sup>72</sup> See eg, Shannon Bond, 'Your Facebook Account Was Hacked. Getting Help May Take Weeks — Or \$299', (NPR, 2 August 2021) <<https://www.npr.org/2021/08/02/1023801277/your-facebook-account-was-hacked-getting-help-may-take-weeks-or-299>> accessed 15 July 2022.

<sup>73</sup> Danielle Keats Citron & Mary Anne Franks, 'Criminalizing Revenge Porn' (2014) 49 Wake Forest Law Review 345.

<sup>74</sup> See eg, Lee Moran, 'Donald Trump's Small Desk is Now Part of a Hilarious 'Photoshop Battle'', *The Hill* (November 28, 2020), <[https://www.huffpost.com/entry/donald-trump-tiny-desk-photoshop-battle\\_n\\_5fc1f423c5b66bb88c670e5f](https://www.huffpost.com/entry/donald-trump-tiny-desk-photoshop-battle_n_5fc1f423c5b66bb88c670e5f)> accessed 15 July 2022.

<sup>75</sup> The 2021 Rules, s 12.

<sup>76</sup> *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, UN Doc A/HRC/38/35 (UNHRC, 2018), <<https://undocs.org/A/HRC/38/35>> accessed 15 July 2022. See also, 'The Social Media Councils: Consultation Paper', (Article 19, 2019) <<https://www.article19.org/wp-content/>

to social media platforms, but to all publishers of news, current affairs and online curated content, comprising a significant subset of all online content.<sup>77</sup> This breadth in scope creates an overlap with existing self-regulatory organizations, such as the Press Council of India and the Advertising Standards Council of India.

While recognition of a role for self-regulatory structures in the social media realm could potentially be a positive development, any actual impact of these bodies on the development of content governance in India is undermined by the fact that the Rules grant an Inter-Departmental Committee organized by the Ministry of Information and Broadcasting with the power to oversee and overrule their decisions.<sup>78</sup> This Inter-Departmental Committee would exercise enormous power over online speech, retaining an ability to reclassify, edit, delete, or otherwise modify content from online news and curated content publishers.<sup>79</sup> In this context, it is difficult to see the purpose of the self-regulatory body at all, especially given that they are wholly dependent on the Inter-Departmental Committee to translate their recommendations into actionable orders.<sup>80</sup> Once again, the inclusion of these nominally independent agencies seems to serve little function other than potentially providing a layer of insulation to shroud the government's assertion of vast new coercive powers over the online political discourse.

## V. CONCLUSION

The lack of adequate representation for Global South interests and perspectives is one of the most significant structural problems in modern content governance.<sup>81</sup> Although the biggest platforms have pivoted away from “marinating in the First Amendment”,<sup>82</sup> and towards a more nuanced understanding of freedom of expression as an international human right,<sup>83</sup> the centre of

---

uploads/2019/06/A19-SMC-Consultation-paper-2019-v05.pdf> accessed 27 December 2021.

<sup>77</sup> The 2021 Rules, s 8.

<sup>78</sup> The 2021 Rules, ss 13-14.

<sup>79</sup> The 2021 Rules, ss 14-16.

<sup>80</sup> The 2021 Rules, s 14(5).

<sup>81</sup> Jenny Domino, ‘Why Facebook’s Oversight Board is Not Diverse Enough’, (*Just Security*, 21 May 2020) <<https://www.justsecurity.org/70301/why-facebooks-oversight-board-is-not-diverse-enough/>> accessed 15 July 2022.

<sup>82</sup> David Kaye, Panelist, ‘Glasnost! Nine Ways Facebook can Make itself a Better Forum for Free Speech and Democracy’ (*Bonaverro Institute of Human Rights*, 28 February 2019), <<https://livestream.com/oxuni/facebook-freespeech-democracy/videos/188101461>> accessed 15 July 2022.

<sup>83</sup> See, eg, ‘The References to Relevant Human Rights Principles in Guiding Decision-making at the Facebook Oversight Board: Oversight Board Bylaws’, (*Facebook*, January 2020)



gravity for these conversations remains heavily focused on the United States and other western democracies.<sup>84</sup> Governments, particularly from emerging democracies, have an important role to play through informal pressure and, where appropriate, through regulation, in pushing platforms to pay more attention to the needs of their international user base. As the largest market for Facebook and YouTube, and one of the largest markets for Twitter, India is well positioned to play a leading role in fostering a pivot by these platforms towards allocating their resources and attention in a manner that corresponds to their global footprint.<sup>85</sup>

Unfortunately, rather than supporting better and more responsive policymaking at online platforms, the 2021 Rules pose a significant challenge to freedom of expression in India, by pushing the platforms into a position where their moderation processes are subject to being further weaponized for political ends. Worse still, by bundling some nominally positive provisions, such as transparency requirements and recognition of a co-regulatory structure, alongside more nakedly repressive changes, the 2021 Rules pettifog the legitimate content governance problems that exist, and discredit real and meaningful efforts at improving accountability. Nationalistic anger at having India's online discourse controlled by foreign entities is fully understandable. But solutions to this challenge which comes at the cost of the very characteristics that made these platforms popular and useful in the first place, namely their free and open character, will ultimately be counterproductive to freedom of expression and democracy in India.

It is also worth noting that, for all the well-documented complaints about content moderation, there can be benefits underlying the platforms' exercise of editorial control over the discourse they host, including where it has been used to push back against harmful government narratives. Their relatively aggressive crackdown on COVID-19 misinformation, including State-driven misinformation, is one example which has been inarguably in the public interest.<sup>86</sup> Likewise, the platforms' pushback against attempts by former U.S.

---

<<https://about.fb.com/wp-content/uploads/2020/01/Bylawsv6.pdf>> accessed 15 July 2022.

<sup>84</sup> Chinmayi Arun, 'Facebook's Faces' 135 *Harvard Law Review Forum*.

<sup>85</sup> 'Leading Countries Based on Facebook Audience Size as of October 2021 (in Millions)', (*Statista*, 2021) <<https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>> accessed 15 July 2022. 'India Becomes YouTube's Largest and Fastest Growing Market', *The Hindu* (9 April 2019) <<https://www.thehindu.com/business/india-becomes-youtubes-largest-and-fastest-growing-market/article26785428.ece>> accessed 15 July 2022; 'Leading Countries Based on Number of Twitter Users as of October 2021 (in millions)', (*Statista*, 2021) <<https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>> accessed 15 July 2022.

<sup>86</sup> See, eg, Kurt Wagner, 'Facebook, Twitter, YouTube Remove Posts From Bolsonaro', *Bloomberg* (30 March 2020) <<https://www.bloomberg.com/news/articles/2020-03-31/>>

President Donald Trump to stir up unrest related to his loss in the 2020 election was clearly to the good of that country's democracy, especially in light of the shocking escalation that followed.<sup>87</sup> While there are legitimate concerns about accountability deficits underlying these decisions, particularly in the context of the Global South, legislative responses which aim to bring the platforms under government control are, in the long run, a greater threat to democracy and freedom of expression.

Instead, India should target its policies towards empowering users and fostering real platform transparency and accountability. This is an opportunity for Indian leadership, as the country has traditionally been a global leader in recognition of core democratic values, such as the right to information.<sup>88</sup> Yet, here too, recent years have seen an erosion of this cornerstone of Indian democracy.<sup>89</sup> There is a clear nexus between the deteriorating health of India's own democratic and constitutional order, and the closing window of opportunity to play a leading role in pushing for a global content governance structure which respects the needs and interests of India's massive online community.

As noted in the introduction, India's government is by no means unique in seeking to weaponize the power of social media platforms to entrench government interests. However, among the broader tide of democratic backsliding, one major asset which India has, and which is missing among many contemporaries across the Global South, is the strength and independence of its judiciary. Although there are some troubling signs that the government

---

facebook-twitter-pull-misleading-posts-from-brazil-s-bolsonaro [<https://perma.cc/YG6E-6WMT>] accessed 15 July 2022]; See Kim Lyons, 'Twitter Removes Tweets by Brazil, Venezuela Presidents for Violating COVID-19 Content Rules', (*The Verge*, 30 March 2020) <<https://www.theverge.com/2020/3/30/21199845/twitter-tweets-brazil-venezuela-presidents-covid-19-coronavirus-jair-bolsonaro-maduro> [<https://perma.cc/S46G-PWLS>]> accessed 15 July 2022; See Rebecca Shabad, 'Twitter Removes Tweet Highlighted by Trump Falsely Claiming COVID-19 'Cure'', (*NBC News*, 28 July 2020) <<https://www.nbcnews.com/politics/donald-trump/twitter-removes-tweet-highlighted-trump-falsely-claiming-covid-cure-n1235075> [<https://perma.cc/U2PP-PBLU>]> accessed 15 July 2022

<sup>87</sup> Megan Graham & Salvador Rodriguez, 'Twitter and Facebook Race to Label a Slew of Posts Making False Election Claims before all Votes Counted', (*CNBC*, 4 November 2020) <<https://www.cnn.com/2020/11/04/twitter-and-facebook-label-trump-posts-claiming-election-stolen.html>> accessed 15 July 2022.

<sup>88</sup> *S.P. Gupta v Union of India* 1981 Supp SCC 87.

<sup>89</sup> Yashovardhan Azad and M Sridhar Acharyulu, 'RTI: A bill that may kill a right', (*Hindustan Times* (22 July 2019) <<https://www.hindustantimes.com/analysis/rTI-a-bill-that-may-kill-a-right/story-EkYRO9yihPgoOexQJaCGJ.html>> accessed 15 July 2022; Anjali Bhardwaj and Amrita Johri, 'To defend Modi govt's RTI Act Changes, BJP Released a 'Factsheet'. It doesn't have Much Facts', (*The Print*, 24 July 2019) <<https://theprint.in/opinion/to-defend-modi-govts-rTI-act-changes-bjp-released-a-factsheet-it-doesnt-have-much-facts/267161/>> accessed 15 July 2022.

is trying to push the envelope in circumventing judicial oversight over their actions, a constitutional challenge is nonetheless the most promising avenue for curtailing the worst aspects of the 2021 Rules.<sup>90</sup> Given the Court's prior role in recognizing the importance of intermediaries in the exercise of a robust right to freedom of expression, and the current raft of legal challenges being advanced against the 2021 Rules,<sup>91</sup> there is hope that that institution will remain as a bulwark in support of India's constitutional values, and against the common tide of platform-enabled authoritarianism sweeping the world.<sup>92</sup>

---

<sup>90</sup> 'I&B Sought Media Compliance with IT Rules Even after HC Issued Stay, Digital Rights Body Says', (*The Wire*, 26 January 2022) <<https://thewire.in/government/ib-sought-media-compliance-with-it-rules-even-after-hc-issued-stay-digital-rights-body-says>> accessed 15 July 2022.

<sup>91</sup> 'Bombay High Court stays the Operation of Rule 9(1) and Rule 9(3) of IT Rules, 2021' (*Internet Freedom Foundation*, 2021) <<https://internetfreedom.in/bombay-high-court-stays-the-operation-of-rule-9-1-and-rule-9-3-of-it-rules-2021/>> accessed 15 July 2022.

<sup>92</sup> 'NH Web Desk, 'Joint Plea by 13 Media Houses Challenges Constitutional Validity of IT Rules, 2021 in Madras High Court' *National Herald* (23 June 2021), <<https://www.nationalheraldindia.com/india/joint-plea-by-13-media-houses-in-madras-hc-challenges-constitutional-validity-of-it-rules-2021>> accessed 15 July 2022.