



**FRIEDRICH NAUMANN
FOUNDATION** For Freedom.

South Asia

STUDENT ESSAY COMPILATION

International Summer School on Platforms and Democracy

*Beyond Digital India Act, Marco Civil, and EU
Digital Services Act*

April 22-28, 2024



LEIBNIZ INSTITUTE
FOR MEDIA RESEARCH
HANS-BREDOW-INSTITUT

Table of Contents

1.Introduction	1
2. How can platforms incorporate systemic accountability within their existing compliance frameworks and take human rights considerations into account?	2
3. The Role of Transparency Reporting in Ensuring Platform Accountability and User Rights Online	5
4. The fragile line between safeguarding elections and violating free speech: Should governments intervene in platform content Moderation?	9
5. Overseeing Meta's Oversight Board	12
6. Grounds and Limits to the Protection of Online Anonymity	15

1.Introduction

Between April 22-28, 2024 the Centre for Communication Governance at National Law University Delhi (CCG NLUD) hosted the Summer School on Platform Governance organised between the Hans-Bredow Institut, University of Hamburg, the Institute for Technology and Society of Rio De Janeiro (ITS Rio), and CCG NLUD. The goal of the week-long program is to bring together a diverse set of scholars, instructors, and students for an intensive week of lectures, discussions, and interactive classroom activities. While past editions of the Summer School have focussed on Data Governance, Freedom of Expression & Online Intermediaries, Privacy, Harmful Speech, Surveillance, and Data Protection, this year, the focus of the Summer School will be on “Platforms and Democracy: Beyond Digital India Act, Marco Civil and EU Digital Services Act”.

The pervasive presence of platforms in our lives, their dynamic nature and their ever-expanding role as facilitators of active civic participation has been widely celebrated. Platforms have empowered participation in the public discourse in ways that were previously not envisioned. More recently, however, platforms have been criticised for having a negative impact on civic participation and democratic processes. Charges ranging from interference in electoral processes to algorithmic promotion of hate speech, have been levelled in multiple countries. Consequently, platforms are currently being examined through several regulatory approaches, including frameworks on media law, competition law, digital public goods etc.

Brazil, Germany, and India are home to vibrant democracies and diverse cultural and linguistic backgrounds. Each country has in the recent past made significant regulatory efforts to regulate platforms, and the Summer School seeks to harness these diverse learnings and facilitate a cross-country exchange of ideas.

The Summer School featured thirty-two students from Brazil, Germany, and India. As part of their involvement in the Summer School, the students were required to prepare short essays on contemporary issues within the realm of platform governance.

The essay topics cover several important themes that come within the ambit of platform governance such as the human rights obligations of platforms, including free speech and due process, anonymity, transparency, and the role of platforms in elections. To this end, students prepared essays on platforms’ human rights obligations; evaluating Meta’s Oversight Board as a tool for accountability; the benefits and risks of anonymity; transparency mechanisms for platform accountability, and the impact of platforms on elections. The following essays were prepared by the students of the Summer School.

2. How can platforms incorporate systemic accountability within their existing compliance frameworks and take human rights considerations into account?

Daniela Leal, Khushi Jaffar and Jonathan Nörz

Communicational human rights and platform governance

Communicational human rights are pivotal to the functioning of democracies. They can be critically assessed by examining the role and responsibility of Social Media Platforms in upholding these rights. Instances such as the attack on Brazil's seat of government coordinated through Big Techs' platforms,¹ or the sudden lockout of a political party's account during a critical election period,² underscore the potential for human rights violations when individuals or groups interact with social media platforms.

Several mechanisms can be incorporated into one platform such as governmental committees or algorithm-based ex-ante content moderation. However, there is a pressing need to examine the efficacy of operationalising fundamental human rights principles within domestic legal instruments seeking to regulate processes pertaining to content moderation.

To accomplish that, this essay will examine select contemporary approaches towards “human rights due diligence” (hereinafter **“HRDD”**) as a component of platform governance efforts in various jurisdictions. It will be framed in the context of the emerging universalist consensus around the issue, as represented in the UNESCO Guidelines for Regulating Digital Platforms for Information as a Public Good (2023) (**“UNESCO Guidelines”**) and the UN Guiding Principles on Business and Human Rights (**“UN Business Principles”**).

¹ Bruno Fonseca and Laura Scofield, “*Bolsonaristas Usam Código “Festa da Selma” para Coordenar Invasão em Brasília*”, Agência Pública, <<https://revistagalileu.globo.com/sociedade/noticia/2023/01/bolsonaristas-usam-codigo-festa-da-selma-para-coordenar-invasao-em-brasilia.ghtml>>; and Carlos Affonso Souza, “*Invasão de Brasília abre debate sobre violência nas redes sociais*.” Universo Online (UOL), <<https://www.uol.com.br/tilt/colunas/carlos-affonso-de-souza/2023/01/09/invasao-de-brasilia-abre-debate-sobre-violencia-nas-redes-sociais.htm>>. (accessed 25 April 2024).

² Cf. press release of the German Federal Constitutional Court <<https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/DE/2019/bvg19-038.html;jsessionid=670E151595AB9DD76FB484A0D26445A7.internet951>> (accessed 5 April 2024), ruling in favor of the reinstallation of the party's account by way of a temporary injunction.

The imperative for human rights due diligence

In this essay, HRDD is viewed as a code of behaviour stemming from platforms' obligation to safeguard human rights or minimize potential risks.³ Instead of considering HRDD as a tool to erase all possible risk, it is rather understood as “*pragmatic standard that strives to minimize risks to human rights [...] without disproportionately limiting individual freedom*”.⁴ HRDD finds mention in Principle 1 of the UNESCO Guidelines as an affirmative obligation.⁵ In the UN Business Principles, there is a key obligation on businesses to formulate outcome-oriented strategies to respect and protect human rights during operations. However, the operationalization of this concept in domestic law often falls short of its intended purpose.

In fact, recent research shows that the adoption of a human rights-centric approach in platform regulation has encountered resistance from certain states. This has resulted in a significant disregard for human rights in formulation of principles governing digital platform governance, with the exception of initiatives such as EU's Digital Services Act (“DSA”) and early United Nations efforts.⁶ As a result, endeavours like those undertaken by UNESCO to establish fundamental standards for governmental oversight become crucial for promoting a global approach that embraces HRDD in a collaborative manner. Nonetheless, criticism directed at UNESCO Guidelines highlight its failure to adequately prioritize human rights, particularly in regions marked by authoritarian regimes or fragile democracies.⁷ They underscore the need for continued efforts towards addressing HRDD considerations.

Notwithstanding these critiques, UNESCO has proactively initiated actions by assembling a network of think tanks and research centres specializing in digital policy globally. This network aims to assess the application of its guidelines across various regulatory environments worldwide.⁸ Indeed, the path toward regulatory implementation poses numerous challenges, especially concerning the incorporation of HRDD by digital platforms. Addressing these

³ Björnstjern Baade, 'Due Diligence and the Duty to Protect Human Rights' in Heike Krieger/Anne Peters/Leonhard Kreuzer (eds.), *Due Diligence in the International Legal Order* (Oxford: OUP 2020) 93–95.

⁴ Id. 108.

⁵ UNESCO, “Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms” (2023) <<https://unesdoc.unesco.org/ark:/48223/pf0000384031.locale=en?ref=static.internetfreedom.in>> (accessed 5 April 2024).

⁶ Yasmin Afina et.al., Towards a Global Approach to Digital Platform Regulation Chatham House (8 January 2024) <<https://www.chathamhouse.org/sites/default/files/2024-01/2024-01-17-towards-global-approach-digital-platform-regulation-afina-et-al.pdf>>. (accessed 5 April 2024).

⁷ Laura et.al., ‘UNESCO’s guidelines for regulating digital platforms must center human rights for everyone, everywhere’ Access Now (19 May 2023) <<https://www.accessnow.org/unescos-guidelines-for-regulating-digital-platforms-must-center-human-rights-for-everyone-everywhere/>>. (accessed 25 April 2024).

⁸ I4T Knowledge, ‘Contribution to the UN Global Digital Compact’, Press Release <https://sites.google.com/i4tknowledge.net/i4tknowledge2024wd/press-release?utm_source=Suscripci%C3%B3n+OBSERVACOM&utm_campaign=d0c044084c-EMAIL_CAMPAIGN_2024_02_29_09_53&utm_medium=email&utm_term=0_-d0c044084c-%5BLIST_EMAIL_ID%5D>. (accessed 5 April 2024).

complexities will be crucial for fostering effective and rights-respecting regulatory frameworks in the digital environment.

Challenges to human rights due diligence: corporate prioritization

For instance, the increasingly corporatized nature of HRDD in the information and communication technology sector is incompatible with the human rights project by design. This is because the corporate purpose of maximizing shareholder value, made effective through the combination of limited liability and separate legal personality, reflects in the execution of the “fiduciary” mandate of companies, incentivizing them away from prioritizing human rights protection over shareholder satisfaction.⁹

The results of Article 19’s Pilot Project Outcome Report show some significant issues. These include: (a) Internet registries and registrars not making clear commitments to human rights; (b) not paying enough attention to how their products affect communities; and (c) not having established human rights standards when working with others. Furthermore, it continues to remain difficult to include the online impact of these companies' main activities in the risk-and-disclosure framework, signaling the need to investigate further.¹⁰

In conclusion, the interconnectedness of social media platforms with human rights violations underscores the urgent need for robust regulatory frameworks. Mechanisms such as HRDD, as outlined in UNESCO Guidelines and UN Business Principles do offer promising avenues for safeguarding rights within digital environments. However, their effective implementation faces obstacles, including resistance from certain states and corporate prioritization of profit over rights. Critiques of current approaches highlight deficiencies, particularly in regions with authoritarian regimes. Yet, proactive efforts, like UNESCO's ongoing evaluation network, demonstrate a commitment to addressing these challenges. Moving forward, bridging the gap between corporate interests and human rights imperatives will be essential for creating truly human rights-respecting digital ecosystems.

⁹ Alex Kyriakakis, 'Corporate Law's Threat to Human Rights: Why Human Rights Due Diligence Might Not Be Enough' (2021) 6(1) *Business and Human Rights Journal* <<https://www.cambridge.org/core/journals/business-and-human-rights-journal/article/corporate-laws-threat-to-human-rights-why-human-rights-due-diligence-might-not-be-enough/CC548D67B5986DE3CE80101598DA3CEC>> (accessed 25 April 2024).

¹⁰ Article 19 and Danish Institute for Human Rights, 'Pilot Project Outcome Report' (2021), <https://www.article19.org/wp-content/uploads/2021/06/A19-and-DIHR-pilot-project-outcome-report_FINAL.pdf> (accessed 25 April 2024).

3. The Role of Transparency Reporting in Ensuring Platform Accountability and User Rights Online

Annabell Günther, Eshita Razdan, Ana Paula Policarpo and Wiesna Ruhin

Transparency is indispensable in a society increasingly shaped by digital platforms to ensure accountability and safeguard users' rights.¹¹ Platforms have employed several self-regulatory transparency measures, including the publication of Community Standards, to inform their users about what content is permissible and the internal procedures that govern content takedowns.¹² Another approach to increase transparency is for platforms to submit voluntarily to audits or assessments conducted by third parties.¹³

Yet, in some countries, such as Germany, transparency measures rely on legal obligations. In accordance with the EU Digital Services Act (DSA), digital platforms must provide data access for independent researchers and publish transparency reports.¹⁴ This essay examines the importance of transparency in holding platforms accountable and safeguarding users' rights. Specifically, it delves into transparency reporting and evaluates its effectiveness in achieving these objectives.

Transparency reports consist of aggregated data and qualitative information about moderation actions, disclosures, and other practices concerning user-generated content and government surveillance.¹⁵ They intend to demonstrate the values of digital platforms as well as their efforts to protect users' rights.¹⁶ However, there is a lack of consistency concerning the criteria, data-counting, and definitions leading to a low reliability of transparency reports.¹⁷

¹¹ Caitlin Vogus/Emma Llansó, "Making Transparency Meaningful: A Framework for Policymakers - Introduction" (2021) 5 (5).

¹² Robert Gorwa/Timothy Garton Ash, "Democratic Transparency in the Platform Society" (2020), 286 (296).

¹³ Robert Gorwa/Timothy Garton Ash, "Democratic Transparency in the Platform Society" (2020), 286 (299).

¹⁴ David Nosák, "The DSA Introduces Important Transparency Obligations for Digital Services, but Key Questions Remain" (2021)

<<https://cdt.org/insights/the-dsa-introduces-important-transparency-obligations-for-digital-services-but-key-questions-remain/>> accessed 25.04.2024.

¹⁵ Caitlin Vogus/Emma Llansó, "Making Transparency Meaningful: A Framework for Policymakers - Introduction" (2021) 5 (6).

¹⁶ Kevin Bankston/Ross Schulman, "Case Study #3: Transparency Reporting" <<https://www.newamerica.org/in-depth/getting-internet-companies-do-right-thing/case-study-3-transparency-reporting/>> accessed 25.04.2024.

¹⁷ Caitlin Vogus/Emma Llansó, "Making Transparency Meaningful: A Framework for Policymakers - Transparency Reports" (2021) 8 (14).

Impact of Platforms on User Rights

Digital platforms serve as both "content guardians," determining what content is permitted or prohibited, and content organizers, prioritizing certain content over others. However, the moderation of inappropriate content and the prioritization of relevant content by these platforms can infringe on freedom of expression by deliberately restricting or expanding the reach of specific information.¹⁸ As such, it is essential for all sectors of society, particularly users affected by content moderation, to closely monitor these platform interventions, and transparency should help users and policymakers make informed choices.

Additionally, transparency in data practices is crucial for protecting users' privacy and ensuring data security. Online platforms collect vast amounts of user data for personalization and targeted advertising, raising concerns about privacy and security. Transparent data practices involve informing users about data collection, usage, and sharing practices.¹⁹

Leveraging Transparency for Informed Policy-Making against Online Harms

Transparency is necessary in combating online harms on digital platforms.²⁰ Through the revelation of content moderation policies, processes, and enforcement measures, intermediaries empower users to understand what constitutes acceptable behaviour and harmful content.²¹ This transparency enables users to make informed decisions about their online interactions and report inappropriate or harmful content effectively.²²

Moreover, transparency fosters accountability among platform operators. When platforms are transparent about their actions and decisions regarding content moderation, they can be held accountable by users, regulators, and civil society organizations.²³ This accountability encourages platforms to invest in more robust moderation systems, respond promptly to reports of harmful content, and continuously improve their policies to mitigate online harms successfully.²⁴

¹⁸ João Carlos CORREIA, "O admirável mundo das notícias". (LabCom Books 2011), 249.

¹⁹ KURTZ, Lahis Pasquali; DO CARMO, Paloma Rocillo Rolim; VIEIRA, Victor Barbieri Rodrigues. 'Transparency in content moderation: national regulatory trends'. (IRIS-BH) <<https://bit.ly/3xjAUka>> accessed 26.04.2024.

²⁰ Department for Digital, Culture, Media & Sport (2019), Online Harms White Paper. Accessible at: https://assets.publishing.service.gov.uk/media/605e60c6e90e07750810b439/Online_Harms_White_Paper_V2.pdf

²¹ Nicolas P. Suzor, Sarah Myers West, Andrew Quodling, Julian York, What do we mean when we talk about transparency? Toward meaningful transparency in commercial moderation. Accessible at: <https://ijoc.org/index.php/ijoc/article/view/9736>

²² The Santa Clara Principles, On Transparency and Accountability in Content Moderation

²³ Robert Gorwa, Timothy Garton Ash, Democratic Transparency in the Platform Society, Accessible at: <https://www.cambridge.org/core/books/social-media-and-democracy/democratic-transparency-in-the-platform-society/F4BC23D2109293FB4A8A6196F66D3E41>

²⁴ Department for Digital, Culture, Media & Sport (2019), Online Harms White Paper. Accessible at: https://assets.publishing.service.gov.uk/media/605e60c6e90e07750810b439/Online_Harms_White_Paper_V2.pdf

Furthermore, transparency builds trust between platforms and their users.²⁵ When users trust that a platform is transparent about its operations and committed to combating online harms, they are more likely to engage in positive online behaviours and support the platform's efforts to maintain a safe and healthy digital environment.²⁶ Trust also encourages users to report instances of harmful content and collaborate with platforms to address emerging threats.²⁷ Overall, transparency plays a pivotal role in empowering users, holding platforms accountable, and fostering a culture of safety and responsibility online.²⁸ Platforms that prioritize transparency are better equipped to mitigate online harms and create a more positive and inclusive online experience for all users.²⁹

Examples of Regulatory Measures and Transparency Initiatives

Recognising the pivotal role of content moderation in online discourse, regulatory directives aim to enhance transparency in platform practices by mandating the publication of reports on content moderation activities. India's IT (Intermediary Guidelines and Digital Media Ethics Code) Rules have a provision mandating Social Media Intermediaries (SSMIs) to release monthly transparency reports concerning their content moderation endeavours.³⁰ While this signifies a positive stride, the heavy reliance on statistics alongside the wide discretion afforded to platforms in framing these reports undermines the essence of meaningful transparency.³¹ Moreover, adherence to the IT Rules does not necessitate disclosure of the number of government content removal requests, which is crucial for achieving true transparency.³²

Similarly, transparency reports under Germany's Network Enforcement Act (NetzDG) exemplify a pioneering step towards platform accountability. These reports, stipulated by law, divulge

²⁵ Department for Digital, Culture, Media & Sport (2019), Online Harms White Paper. Accessible at: https://assets.publishing.service.gov.uk/media/605e60c6e90e07750810b439/Online_Harms_White_Paper_V2.pdf

²⁶ David Nosák, The DSA Introduces Important Transparency Obligations for Digital Services, but Key Questions Remain, Accessible at: <https://cdt.org/insights/the-dsa-introduces-important-transparency-obligations-for-digital-services-but-key-questions-remain/>

²⁷ Shagun Jhaver, Do users want platform moderation or individual control? Examining the role of third-person effects and free speech support in shaping moderation preferences, Accessible at: <https://journals.sagepub.com/doi/abs/10.1177/14614448231217993>

²⁸ David Nosák, The DSA Introduces Important Transparency Obligations for Digital Services, but Key Questions Remain, Accessible at: <https://cdt.org/insights/the-dsa-introduces-important-transparency-obligations-for-digital-services-but-key-questions-remain/>

²⁹ Center for Technology and Society, Platform Transparency Reports - Just how transparent?, Accessible at: <https://www.adl.org/resources/blog/platform-transparency-reports-just-how-transparent>

³⁰ Vasudev Devadasan, 'Why Transparency Reporting under the IT Rules 2021 Is a Mess' (MediaNama, 19 April 2022) <https://www.medianama.com/2022/04/223-transparency-reports-social-media-platforms-unhelpful/> accessed 26 April 2024.

³¹ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Rule 4(1)(d).

³² Anandita Mishra, '#SocialMediaComplianceWatch: Analysis of Social Media Compliance Reports for the Month of February 2022' (Internet Freedom Foundation, 26 May 2022) <https://internetfreedom.in/compliance-report-feb-2022/> accessed 26 April 2024.

figures pertaining to flagged posts and subsequent removals, setting a global precedent.³³ However, a holistic approach to transparency should encompass procedural transparency beyond numerical representation.

On the other hand, Brazilian legislation, Marco Civil da Internet,³⁴ does not establish the creation and dissemination of transparency reports by platforms, but it does provide for user rights, including the right to information.

Emphasising process transparency could entail divulging metrics like the number of in-house and third-party content moderators within social media companies, along with diverse language capabilities to ensure broader representation in human content moderation.³⁵ The significance of transparency reports lies not merely in figures, but in elucidating the adjudication and removal processes. Numbers, devoid of contextual insights, offer limited utility; therefore, discussions on developing new mechanisms to enhance understanding of platform content moderation for broader transparency, along with measures to improve researcher access,³⁶ are gaining momentum.

In conclusion, transparency is a fundamental pillar of accountability and user protection in the digital age. By embracing transparency measures, digital platforms can cultivate a safer and more inclusive online environment, where user rights are respected, and harmful content is effectively addressed.

³³ Heidi Tworek, 'How Transparency Reporting Could Incentivize Irresponsible Content Moderation' (Centre for International Governance Innovation, 10 December 2019) <https://www.cigionline.org/articles/how-transparency-reporting-could-incentivize-irresponsible-content-moderation/> accessed 26 April 2024.

³⁴ 'L12965' (PRESIDÊNCIA DA REPÚBLICA, 23 April 2014) <www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm> accessed 26 April 2024.

³⁵ Lab Platform Governance Media and Technology (PGMT), 'Platforms Overwhelmingly Use Automated Content Moderation, First DSA Transparency Reports Show' (Lab Platform Governance Media and Technology PGMT) <https://platform-governance.org/2023/platforms-overwhelmingly-use-automated-content-moderation-first-dsa-transparency-reports-show/> accessed 26 April 2024.

³⁶ Daphne Keller, 'Some Practical Postulates about Platform Data' (Center for Internet and Society) <https://cyberlaw.stanford.edu/blog/2023/05/some-practical-postulates-about-platform-data> accessed 26 April 2024.

4. The fragile line between safeguarding elections and violating free speech: Should governments intervene in platform content moderation?

Amanda Chami, Pranika Goel, Sakshi Mohan and Julius Tiedmann

As social media platforms establish their primacy over traditional news outlets, they have become prolific grounds for disseminating misinformation. From fabricated allegations to adulterated videos depicting political opponents, misinformation proliferates unchecked, casting doubt on the integrity of the electoral process. In light of this fake news epidemic, how should democratic societies navigate the delicate balance between safeguarding free speech and curbing the harmful effects of false information during critical electoral periods? This essay delves into the advantages and drawbacks surrounding government intervention in combating misinformation on social media platforms during election times.

Arguments against government regulation:

1. The Right to Free Speech

The freedom of speech and expression is one of the most sacrosanct rights in a democratic society. It even protects individuals or groups with whom we fundamentally disagree. This freedom is not just a mere right, but also a limit on the power of government to shrink the space for constitutionally protected speech.

Firstly, misinformation is context-dependent, therefore what the government terms as misinformation may just be protected speech.

Secondly, the government may misuse its power to censor protected speech. As an interested party, the government can lobby platforms to suppress content or add context, leading to biased reporting. Making the government an arbitrator of misinformation can result in censorship, which would involve the threat of retaliation or coercion against online platforms if they do not toe the government lines.

Lastly, misrepresentation is not specified as a ground in the constitution/statute of major countries such as India and the United States as a reasonable restriction on free speech. The government cannot create new grounds beyond the constitutional limits.

2. Preemptive censorship may stifle public discourse

Platform regulation in the past has shown that there is a threat of censoring content preemptively. For instance, the German Network Enforcement Act (NetzDG) has led to an increase in blocking and deletions. 20% of reports were likely justified, as they concerned potentially unlawful

content. However, almost 80% of all reports resulted in no subsequent action, indicating that the bulk of these reports were without merit.³⁷

If platforms cannot accurately identify content like hate-speech or misinformation, companies could be more likely to delete too much rather than too little, if the regulation foresees sanctions.³⁸ If this risk comes true, this will not only undermine the user's rights to free speech, but also the platform's role as a space for the democratic debate.

3. Places disproportionate burden on smaller companies

Moreover, content moderation at an adequate scale is only feasible for large tech companies. Smaller platforms or platforms with fewer resources lack the capability to effectively moderate content, increasing the possibility of excessive content restriction. Therefore, an imposed duty to moderate all content posted online would place a disproportionate burden on smaller companies. Of course, this aspect could be mitigated if the burden imposed by legislation is proportionate to the size of each platform.

Arguments for government intervention:

1. The right to free speech is not absolute

Even though free speech stands as a cornerstone within democratic societies, it could be argued that this right is not absolute. This means that it may allow for mitigation under specific circumstances, particularly when its exercise threatens the integrity of other constitutional guarantees such as the right to free and fair elections, and the preservation of democratic institutions. Since the spread of misinformation can manipulate public opinion and distort electoral outcomes, one may argue that the dissemination of misinformation fails to meet the threshold of "meaningful public discourse" and, therefore, does not merit constitutional protection by governments.

2. Social media shapes democratic processes

Social media possesses distinct features that make it easier to generate and spread fake information. These include features such as personalised content, creation of "echo chambers", and the fact that the business of social media rests upon maximum attention to advertisements, which propels content online to be sensational and thus dangerous. These effects are amplified during elections, where social media can be used as a tool to alter voter preferences, by spreading fake information, shunning off voters from alternative viewpoints and micro-targeting users.

3. Ensuring Accountability of Platforms

The concept of intermediary liability entails that social media platforms are not liable for third-party content they host unless they fail to exercise minimum due diligence. This grants a certain immunity to platforms, which, in many jurisdictions, is not the case for other means of

³⁷ Kasakowskij et. al., Network enforcement as denunciation endorsement? A critical study on legal enforcement in social media, *Telematics and Informatics* 2020, 101317, p. 10.

³⁸ Ceffinato, Hate Speech zwischen Ehrverletzungsdelikten und Meinungsfreiheit, *JuS* 2020, p. 469.

communication, such as radio and television, which tend to be more closely regulated. Extending regulation of fake news to social media platforms during elections is thus an important catalyst in ensuring that platforms take accountability for their role in shaping the public discourse.

Conclusion

The essay highlights the pertinent challenges that the fake news phenomenon poses to the electoral process, indicating that the question will not be *whether* to regulate, but *how*. An ideal regulation would be one that maintains the subtle balance between safeguarding democratic elections and the risk of undermining the democratic debate. Regulations such as the Digital Services Act are already in place and scholars have cited self-regulation by platforms and “light touch” government regulation as viable means of countering fake news. Even though it is only through future research that society will learn how effective or harmful such regulations can really prove to be, we believe that the crucial role social media platforms play in shaping the public debate is reason enough to hold them responsible in cases of online dissemination of misinformation, especially in electoral contexts.

5. Overseeing Meta's Oversight Board

Alice Lana, Karina Chawla, Leon Paul Zgodda and Nandita Yadav

The Oversight Board (hereinafter '*Meta's Oversight Board or OB*') is a body created under Meta's auspices to make precedent-setting content moderation decisions on Facebook, Instagram, and Threads.³⁹ According to the Board's official website, it *"reviews content decisions made by Meta to see if the company acted in line with its policies, values, and human rights commitments"*.⁴⁰ Its decisions are binding, although it can also make non-binding policy recommendations to the company.

Mark Zuckerberg, Meta's founder and CEO, originally described the Board as a kind of *"Supreme Court"*, saying it was a *"new way for people to appeal content decisions to an independent body, whose decisions would be transparent and binding"*.⁴¹ Since October 2020, when it officially began its work, 100 decisions (out of over two million cases that were appealed),⁴² over 250 policy recommendations and 4 advisory opinions have been issued by the Board.

The OB is undeniably a novel approach to platform governance. Among praises and critics, in its 4 years of work, it has shown some of its strengths and limitations. In line with what is pointed out by Wong and Floridi, its main weakness derives from (i) its limited jurisdiction, (ii) lack of diversity, and (iii) lack of control over its precedents. In this short essay, based on the state-of-art discussion, we would like to address these issues and, in the conclusion, some alternatives to the Board's current structure.⁴³

1. Limited Jurisdiction

The OB has been lauded for its authority to determine and clarify its jurisdiction assertively. However, several weaknesses are attached to the limited jurisdiction of the OB. Firstly, the binary decision-making authority of the OB is restrictive. The OB can only give a decision to either affirm or overrule Meta's original content moderation decision. This binary approach⁴⁴

³⁹ 'Oversight Board | Improving How Meta Treats People and Communities around the World' <<https://www.oversightboard.com/>>.

⁴⁰ 'Decisions & Policy Advisory Opinions | Oversight Board' <<https://www.oversightboard.com/decision/>>.

⁴¹ Evelyn Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (2019) 21 North Carolina Journal of Law & Technology 1 <<https://scholarship.law.unc.edu/ncjolt/vol21/iss1/2>>.

⁴² Steven Levy, 'Inside Meta's Oversight Board: 2 Years of Pushing Limits' Wired <<https://www.wired.com/story/inside-metas-oversight-board-two-years-of-pushing-limits/>>.

⁴³ David Wong and Luciano Floridi, 'Meta's Oversight Board: A Review and Critical Assessment' (22 October 2022) <<https://papers.ssrn.com/abstract=4255817>>.

⁴⁴ *ibid.*

prevents the OB from delving into alternative content moderation remedies that could be applied to ambiguous cases. Secondly, the OB has the authority to issue a decision on content moderation pertaining to individual posts and not accounts. This restriction inhibits the OB from dealing with questions that relate to the freedom of expression of individual accounts and not just particular posts.

2. Lack of Diversity

The lack of diversity in the composition of the OB (currently composed of 22 people)⁴⁵ has been highlighted as one of OB's major issues since the backgrounds of people forming the Board also impact the decisions.⁴⁶ It has been evident by the disparity in the representation of Global South countries in the OB, for example, especially considering the number of users from the region. Further, the low representation of the interests of Indigenous communities and LGBTQ+ communities also affects how the board functions to hear and solve disputes. The problem of disparity in representation also creates a chain reaction as the OB's co-chairs and members of the Board nominate its future members and successors. Policy changes relating to diversity must consider questions such as: how can the membership include the views of marginalized and underrepresented groups and how can the Global South be made a more significant part of the Board's decision-making?

3. Lack of Control Over its Precedents

As previously stated, decisions by the OB also serve as a precedent for future cases, enabling a system of case law like in the U.S. judicial process. However, it is not the OB but Meta itself deciding whether to apply a precedent to a new case. Meta has stated that it will enforce the decision for “identical content with a parallel context where technically and operationally feasible”.⁴⁷ This means that Meta will decide whether the set precedent shall govern identical content in a different context or similar content. Thus, if Meta was not sharing the decision of a previous ruling of the OB, it could theoretically decide not to apply that precedent to another case, giving it the power of secretly undermining the OB and its effectiveness in making systematic changes to its content moderation practices.

4. Conclusion:

⁴⁵ ‘Meet The Board | Oversight Board’ (29 April 2024) <<https://www.oversightboard.com/meet-the-board/>>.

⁴⁶ David Wong and Luciano Floridi, ‘Meta’s Oversight Board: A Review and Critical Assessment’ (22 October 2022) <<https://papers.ssrn.com/abstract=4255817>>.

⁴⁷ Monika Bickert, ‘Responding to the Oversight Board’s First Decisions’ (*Meta*, 28 January 2021) <<https://about.fb.com/news/2021/01/responding-to-the-oversight-boards-first-decisions/>> accessed 5 June 2024.

Even though other limitations to the OB's efficiency exist, such as doubts regarding its independence, these three issues showcase some of the Board's limitations. To conclude, we would like to present brief alternatives to tackle these problems. The limited jurisdiction issues could be improved by allowing the OB to issue more nuanced decisions and cover accounts. The lack of diversity could be tackled with affirmative action measures to ensure marginalized populations are sufficiently represented. Finally, one option to counter the lack of control over its precedents is to increase the scope of the OB itself, giving users the right to appeal a content moderation decision because a precedent has been incorrectly or not applied at all.

6. Grounds and Limits to the Protection of Online Anonymity

Akshay Dhekane, Ludmila Lago, Marrit Krebs and Shashank Mehrotra

The digitalization of modern life resulted in an inversion between the proportion of public and private life. In fact, there are currently very few spaces that remain truly private, whereas a considerable amount of personal information is publicly available, regardless of authorization from and beyond the control of their owners.⁴⁸ The extensive publicity includes not having a heavily reduced expectation of secrecy on personal communications, affecting therefore not only privacy, but the right to freedom of speech as well. In fact, if our communications are constantly being monitored by private and public entities,⁴⁹ are we effectively communicating freely or only saying what does not present a risk of retaliation?⁵⁰

In light of this, there are currently many discussions on the right to online anonymity. Those in favour of online anonymity defend it as a remedy to fight the extensive control over online speech, especially because: (i) it helps users to find information and support in topics which are seen as indecorous;⁵¹ (ii) it empowers minority groups to try out different identities while protecting them against possible harassment and stigmatisation;⁵² (iii) it encourages the users to

⁴⁸ As commented by Sergio Branco in its book "Memória e Esquecimento na Internet": "O surgimento das redes sociais acabou acarretando um debate que, se não era inédito, era muito menos relevante para as pessoas comuns – qual o limite entre o público e o privado em suas vidas? [...] Convém lembrar, por fim, que até o término do século 20, somente pessoas que gozavam de prestígio social ou que desempenhavam um papel público relevante tinham suas vidas convertidas em bens acessíveis ao público – quer pela edição de seus diários ou autobiografia ou, ainda, pela produção de obra audiovisual baseada em sua vida. O anonimato e esquecimento eram o destino das pessoas comuns. Já não é mais assim". BRANCO, Sérgio. *Memória e Esquecimento na Internet*. Porto Alegre: Arquipélago Editorial, 2017, p. 47. Free translation: "The advent of social networks ended up leading to a debate that, if not unprecedented, was much less relevant to ordinary people – what is the limit between public and private in their lives? [...] It is worth remembering that until the end of the 20th century, only people who enjoyed social prestige or who played a relevant public role had their lives converted into assets accessible to the public – either by editing their diaries or autobiography or, even, by producing an audiovisual work based on his life. Anonymity and oblivion were the lot of ordinary people. It's not like that anymore".

⁴⁹ Two journalistic papers evidence that Jair Bolsonaro's government has spent more than BRL 31MM in spy software, to illegally monitor private communications of several individuals: MELLO, Patrícia Campos. *Despesas com sistema espião dispararam em estados durante Governo Bolsonaro*. [s.l.], 8 nov. 2022. Available at <https://www.em.com.br/app/noticia/politica/2022/11/08/interna_politica,1418485/despesas-com-sistema-espiao-disparam-em-estados-durante-governo-bolsonaro.shtml>. Access at: 03 apr. 2024; PAES, Caio de Freitas. *Abin de Ramagem gastou R\$31 milhões com ferramentas de vigilância secretas e sem licitação*. [s.l.], 25 abr. 2023. Available at: <<https://apublica.org/2023/04/abin-de-ramagem-gastou-r-31-milhoes-com-ferramentas-de-vigilancia-secretas-e-sem-licitacao/>>. Access at: 15 apr. 2024.

⁵⁰ Studies reveal that people tend to feel more comfortable saying what they want to strangers in an elevator or Uber, than to members of their own communities. For this reason, anonymity may be a relevant tool to foster individuals to communicate freely, without the fear of being rejected. In this sense, please see: RUBIN, Zick, *Disclosing oneself to a stranger: reciprocity and its limits*. *Journal of Experimental Social Psychology*, v. 11, n. 3, p. 233–260, 1975.

⁵¹ 'Right to Online Anonymity' ARTICLE 19 (June 2015);

⁵² Anmol Jain, 'Anonymous Social Media Users: Exploring Policy Solutions to Regulate Anonymity over the Internet' Part II; Lorenzo Ancona, Gabriel Karl, Arnau Marti, Wiktor Samek, 'Has the time come to end anonymity on social media?' *SciencesPo* (2023) P.12

speak out without fear of retaliation; (iv) it provides control over personal information.⁵³ It is worth stressing that, for some groups, online anonymity is not only beneficial but crucial. Only via anonymous accounts can activists and whistleblowers share challenges without the fear of reprisal or discrimination.⁵⁴ To participate in unpopular, marginalised and critical topics is much easier when online anonymity provides confidentiality over identity.⁵⁵

However, while some defend anonymity as a right, for others, the prerogative of using the Internet without having personal information processed would increase a range of issues, including online defamation and hate speech. According to this position, anonymous users cannot be held liable, since they cannot be identified by authorities as the authors of a certain infringing content. As a result, they do not feel the need to be politically correct or inhibit their bias. It is true that absolute anonymity could have negative consequences. For instance, authorities could not identify the author of a certain content deemed as infringing. Yet, unilaterally blaming anonymity for online unlawful conduct is reflective of a limited perspective, for two main reasons, as explained below.

First, there is no such absolute anonymity where authorities cannot identify a user under any circumstance. The way the Internet currently functions makes it practically impossible to navigate it without leaving any traces. Therefore, in practical terms, there are actually different levels of online anonymity, where more or less data is collected.⁵⁶ For instance, a user can refrain from identifying themselves with their real name on a social media platform, but, generally, the platform will still be able to identify the IP address where such user is located.⁵⁷ Thus, online anonymity does not necessarily imply a complete absence of accountability.⁵⁸

Second, despite claims of enhancing security, the imposition of user identification mechanisms increases personal data collection. In turn, the more data is collected by a platform, higher are the chances of it being used by cybercriminals for identity thefts, among other fraudulent mechanisms, not to mention authoritarian governments that may track user data to control political speech online. Hence, not allowing anonymity under the argument that it would promote

⁵³ Rishab Bailey, Vrinda Bhandari, Faiza Rahman 'Examining the Online Anonymity Debate: How Far Should the Law Go in Mandating User Identification?' Data Governance Network, Working Paper 18, 2021 (2022). P.17; Lorenzo Ancona, Gabriel Karl, Arnau Marti, Wiktor Samek, 'Has the time come to end anonymity on social media?' SciencesPo (2023) P.12.

⁵⁴ Jillian York, 'The Right to Anonymity is Vital to Free Expression: Now and Always' EFF (25 March 2020); Anmol Jain, 'Anonymous Social Media Users: Exploring Policy Solutions to Regulate Anonymity over the Internet' Part II; Rishab Bailey, Vrinda Bhandari, Faiza Rahman 'Examining the Online Anonymity Debate: How Far Should the Law Go in Mandating User Identification?' Data Governance Network, Working Paper 18, 2021 (2022) P.8.

⁵⁵ *ibid*

⁵⁶ About the different levels of online anonymity, please check: FLINN, Bill; MAURER, Hermann. Levels of Anonymity. Journal of Universal Computer Science, Graz, v. 1, n° 1, 1995, p. 35-47. Disponível em: <https://www.jucs.org/jucs_1_1/levels_of_anonymity/Flinn_B.pdf>. Acesso em: 03 dez. 2023.

⁵⁷ We say 'generally' since we cannot disregard the possibility of users applying technical mechanisms to hide their real IP addresses, such as VPN connections.

⁵⁸ In this sense: ARANHA, Diego. O que é criptografia fim a fim e o que devemos fazer a respeito? Thomson Reuters: Revista dos Tribunais, 2019. In: DONEDA, Danilo; MACHADO, Diego (coord). A criptografia no direito brasileiro. São Paulo: Thomson Reuters Brasil, 2019, p. 34.

more online security is a paradox, as extensive user identification also exposes individuals to a range of crimes and goes against data minimization principles.⁵⁹

Considering the above, any proposed regulation on the topic should adopt a balanced approach: it should not allow an absolute right to anonymity, where users cannot be identified in any circumstances. At the same time, it should not create extensive identification mechanisms. Internet users should be identified on an exceptional basis, where there is clear public interest in such identification, for instance, for a criminal investigation. This would be possible, for instance, by allowing individuals to use social networks without having to identify themselves with their real names, or an ID number, but requiring platforms to collect minimum data, such as IP addresses.

⁵⁹ Data minimization principles are provided in different legislations, including Article 6, III of Brazilian General Data Protection Law (LGPD) and Article 5(1)(c) of the European Union's General Data Protection Regulation (GDPR).

ABOUT THE NATIONAL LAW UNIVERSITY DELHI (NLUD)

The National Law University Delhi is one of the leading law universities in the capital city of India. Established in 2008 (by Act. No. 1 of 2009), the University is ranked second in the National Institutional Ranking Framework for the last five years. Dynamic in vision and robust in commitment, the University has shown terrific promise to become a world class-institution in a very short span of time. It follows a mandate to transform and redefine the process of legal education. The primary mission of the University is to create lawyers who will be professionally competent, technically sound, and socially relevant, and will not only enter the Bar and the Bench but also be equipped to address the imperatives of the new millennium and uphold constitutional values. The University aims to evolve and impart comprehensive inter-disciplinary legal education which will promote legal and ethical values, while fostering the rule of law.

The University offers a five year integrated B.A., LL.B (Hons.) and one-year postgraduate masters in law (LL.M.) along with professional programs, diplomas and certificate courses for both lawyers and non-lawyers. The University has made tremendous contributions to public discourse on law through pedagogy and research. Over the last decade, the University has established many specialised research centres including the Centre for Communication Governance (CCG), Centre for Innovation, Intellectual Property and Competition, Centre for Corporate Law and Governance, Centre for Criminology and Victimology, and Project 39A. The University has made submissions, recommendations, and worked in advisory/consultant capacities with government entities, universities in India and abroad, think tanks, private sector organisations, and international organisations. The University works in collaboration with other international universities on various projects and has established MoU's with several other academic institutions.

ABOUT THE CENTRE FOR COMMUNICATION GOVERNANCE

The Centre for Communication Governance at the National Law University Delhi (CCG) was established in 2013 to ensure that Indian legal education establishments engage more meaningfully with information technology law and policy and contribute to improved governance and policy making. CCG is the only academic research centre dedicated to undertaking rigorous academic research in India on information technology law and policy in India and in a short span of time has become a leading institution in Asia. Through its academic and policy research, CCG engages meaningfully with policy making in India by participating in public consultations, contributing to parliamentary committees and other consultation groups, and holding seminars, courses and workshops for capacity building of different stakeholders in the technology law and policy domain. CCG has built an extensive network and works with a range of international academic institutions and policy organisations. These include the United Nations Development Programme, Law Commission of India, NITI Aayog, various Indian government ministries and regulators, International Telecommunications Union, UNGA WSIS, Paris Call, Berkman Klein Center for Internet and Society at Harvard University, the Center for Internet and Society at Stanford University, Columbia University's Global Freedom of Expression and Information Jurisprudence Project, the Hans Bredow Institute at the University of Hamburg, the Programme in Comparative Media Law and Policy at the University of Oxford, the Annenberg School for Communication at the University of Pennsylvania, the Singapore Management University's Centre for AI and Data Governance, and the Tech Policy Design Centre at the Australian National University.

The Centre has had multiple publications over the years including reports on Intermediary Liability in India, a report Mapping the Blockchain Ecosystem in India and Australia, an authored UNDP Guide on Drafting Data Protection Legislation, a book on Privacy and the Indian Supreme Court, Hate Speech Report, and most recently two essay series, one on Democracy in the Shadow of Big and Emerging Tech, and a second on Emerging Trends in Data Governance. The Centre has launched freely accessible online databases - Privacy Law Library (PLL) and High Court Tracker (HCT) to track privacy jurisprudence across the country and more than

sixteen jurisdictions across the globe in order to help researchers and other interested stakeholders learn more about privacy regulation and case law. CCG also has an online 'Teaching and Learning Resource' database for sharing research oriented reading references on information technology law and policy. In recent times, the Centre has also offered courses on AI Law and Policy, Technology and Policy, and First Principles of Cybersecurity. These databases and courses are designed to help students, professionals, and academicians build capacity and ensure their nuanced engagement with the dynamic space of existing and emerging technology and cyberspace, their implications for society, and their regulation. Additionally, CCG organises an annual International Summer School in collaboration with the Hans Bredow Institute and the Faculty of Law at the University of Hamburg in collaboration with the UNESCO Chair on Freedom of Communication at the University of Hamburg, Institute for Technology and Society of Rio de Janeiro (ITS Rio) and the Global Network of Internet and Society Research on contemporary issues of information law and policy.



Centre for Communication Governance at National
Law University Delhi, Sector 14, Dwarka, New Delhi,
110078, India

ccgdelhi.org |  @CCGNLUD

Email: ccg@nludelhi.ac

