

# KISIP 2024

Konferensi Ilmu Sosial dan Ilmu Politik



CSIS  
INDONESIA

Google

Research Paper

## Blunt Instruments for Complex Problems: Challenges in Indonesia's Regulation of Disinformation and Content Moderation

---

*Panel 4*

Law, Regulations, and Governance Against Disinformation and on Content Moderation in Indonesia

# Irene Poetranto

**Doctoral Candidate, Department of Political Science,  
University of Toronto**

✉ [irene.poetranto@utoronto.ca](mailto:irene.poetranto@utoronto.ca)

Irene Poetranto is a Doctoral Candidate in the Department of Political Science at the University of Toronto. Her research interests include Internet governance and the politics of Internet regulation. Southeast Asia is her geographic area of focus. Irene is also a Senior Researcher for the Citizen Lab, an interdisciplinary research lab on digital technology and human rights at the Munk School of Global Affairs & Public Policy, University of Toronto.

---

*This working paper is circulated for discussion and comment purposes. It has not been peer-reviewed or been subject to internal review by SAIL, CSIS, or Google. The views expressed here are solely those of the author(s) and do not represent an official position of SAIL, CSIS, Google, or any other organization. Feedback is welcome as the author(s) continue to develop these ideas for future formal publication. Please contact the author(s) directly with any comments or questions.*

Editor: Dandy Rafitrandi

## Abstract

This paper examines the discourse and strategies against the proliferation of false information online (or disinformation) in Indonesia. We explore the role and impact of the Information and Electronic Transactions (ITE) law, the revised Criminal Code, and the Ministry of Communication and Informatics' Ministerial Regulation no. 5/2020, using secondary research and a socio-legal approach. We also compare Indonesia's policies on disinformation against those adopted by other jurisdictions, for example, Germany's Network Enforcement Act (NetzDG). Our findings suggest that efforts to combat disinformation in Indonesia have thus far been ineffective, as they have disproportionately been used against claims of libel or defamation instead. Furthermore, Indonesia's approaches to regulating content are lacking in terms of independent oversight, accountability, and remedy mechanisms. This paper highlights the urgency to reform the relevant laws and policies in Indonesia to better protect against disinformation, improve content moderation, and reduce threats to democratic values and human rights.

**Keywords:** *Social Media, Disinformation, UU ITE, EIT Law, NetzDG*

## Introduction

Concerns regarding disinformation in Indonesia have been raised since the 2014 general election when false information circulated online regarding the presidential candidates. Joko Widodo (popularly known as “Jokowi”), who became president in 2014, was targeted by a disinformation campaign carried out by the supporters of his opponent, Prabowo Subianto, which falsely claimed that Jokowi was of Chinese descent and a Christian.<sup>1</sup> Meanwhile, Jokowi’s supporters also engaged in disinformation, as they circulated false stories about Prabowo.<sup>2</sup> Since then, terms such as “hoaxes,” “fake news,” and “disinformation” have frequently been used in news articles and reports in Indonesia to refer to false information that circulates online. Although Indonesian lawmakers and government officials, including those with the Ministry of Communication and Informatics (MOCI) (*Kementerian Komunikasi dan Informatika*, or Kominfo for short), which is the Internet regulator, have used these terms many times, the government has not released an official definition.

The scholarly literature on how to define and operationalize terms such as fake news and disinformation also continue to remain “fragmentary” and lack “a common vocabulary.”<sup>3</sup> ReRecting this “definitional uncertainty,” states and regional institutions have adopted different approaches to govern false news in the online sphere.<sup>4</sup> For the purposes of this paper on regulating disinformation in Indonesia, the term “disinformation” is used to refer broadly to “false information that is conveyed (intentionally) to confuse other people,” which is a definition offered by the The Large Dictionary of the Indonesian Language (*Kamus Besar Bahasa Indonesia*), the official dictionary of the Indonesian language.<sup>5</sup>

The proliferation of disinformation has been facilitated by the rapid uptake of social media, particularly in highly mobile markets like Indonesia, where people spend over 1ve hours per day on their mobile devices.<sup>6</sup> More than half of Indonesia’s 260 million population were active users of social media in 2019,<sup>7</sup> with the top four most popular

---

<sup>1</sup> CNN Indonesia, “La Nyalla Ungkap Intrik Sebar Hoaks soal Jokowi di 2014,” December 13, 2018, <https://www.cnnindonesia.com/nasional/20181212213201-32-353262/la-nyalla-ungkap-intrik-sebar-hoaks-soal-jokowi-di-2014>.

<sup>2</sup> Pratiwi Utami, “Hoax in Modern Politics: The Meaning of Hoax in Indonesian Politics and Democracy,” *Jurnal Ilmu Sosial Dan Ilmu Politik* 22, no. 2 (January 4, 2019): 86, <https://doi.org/10.22146/jsp.34614>

<sup>3</sup> Samantha Bradshaw et al., “Sourcing and Automation of Political News and Information over Social Media in the United States, 2016–2018,” *Political Communication* 37, no. 2 (March 3, 2020): 174, <https://doi.org/10.1080/10584609.2019.1663322>.

<sup>4</sup> Pattamon Anansaringkarn and Ric Neo, “How Can State Regulations over the Online Sphere Continue to Respect the Freedom of Expression? A Case Study of Contemporary ‘Fake News’ Regulations in Thailand,” *Information & Communications Technology Law* 30, no. 3 (September 2, 2021): 286, <https://doi.org/10.1080/13600834.2020.1857789>

<sup>5</sup> KBBI, “Hasil Pencarian ‘Disinformasi’ - KBBI VI Daring,” *Kamus Besar Bahasa Indonesia (KBBI)*, 2016, <https://kbbi.kemdikbud.go.id/entri/disinformasi>.

<sup>6</sup> Vincent Zhou, “State of Mobile Market 2022: Spotlight on Indonesia Report | Data.Ai Blog,” *Data.Ai* (blog), March 31, 2022, <https://www.data.ai/en/insights/market-data/indonesia-mobile-market-spotlight-2022/>.

<sup>7</sup> We Are Social and Hootsuite, “Digital 2019: Indonesia” (We Are Social and Hootsuite, January 31, 2019), <https://datareportal.com/reports/digital-2019-indonesia>.

platforms being YouTube, WhatsApp, Facebook, and Instagram.<sup>8</sup> Social media has also become a crucial tool for election campaigning, as it is more cost effective for political advertising purposes compared to traditional media (e.g., newspapers or television), and allows those running for office to connect directly with the public.<sup>9</sup>

Social media played a key role in Indonesia's 2019 general election, particularly as it was the first time that presidential and legislative elections were held on the same day.<sup>10</sup> The sheer size and complexity of this election, coupled with the substantial size of voters aged seventeen to thirty-five years old (many of whom were voting for the first time), contributed to the heavy use of social media by political parties to connect with their young, tech-savvy constituents.<sup>11</sup> In the 2024 general election, the number of young or 1st time voters is expected to increase, particularly as millennials and 'generation Z' (or 'gen Z')—defined as those born between 1980 and 1994, and between 1995 and the 2000s, respectively—will make up over fifty-percent of voters in Indonesia.<sup>12</sup>

The public's increased reliance on the Internet and social media to receive and impart information, particularly about elections, has contributed to the circulation of massive amounts of false information online. Disinformation has targeted not only election candidates but also state institutions. For example, the Elections Commission (*Komisi Pemilihan Umum* (KPU)) reported the proliferation of false or misleading content that attacked their role and integrity leading up to the 2019 election, which was unlike previous elections.<sup>13</sup> Consequently, the Indonesian government has become especially concerned with identifying and combating disinformation.

The Indonesian government has blocked online content containing disinformation, which it argued had increased from year to year. In October 2023, Kominfo stated that it had discovered ten times the amount of content containing disinformation since January 2023, compared to the entire year of 2022, and that most of them, or 455 out of 526, were found on Facebook.<sup>14</sup> In response, Kominfo announced in early November 2023 that it had

---

<sup>8</sup> IDN Financials, "Indonesia's Population Up by 3 Million in A Year," *IDN Financials*, May 26, 2021, <https://www.idnfinancials.com/insights/indonesia-s-population-up-by-3-million-in-a-year>.

<sup>9</sup> Laura Paatelainen, Elisa Kannasto, and Pekka Isotalus, "Functions of Hybrid Media: How Parties and Their Leaders Use Traditional Media in Their Social Media Campaign Communication," *Frontiers in Communication* 6 (2022): 1, <https://www.frontiersin.org/articles/10.3389/fcomm.2021.817285>

<sup>10</sup> Jennifer Yang Hui, "Social Media and the 2019 Indonesian Elections: Hoax Takes the Centre Stage," *Southeast Asian Affairs*, 2020, 158, <http://www.jstor.org/stable/26938889>.

<sup>11</sup> Tim Mann, "Six Things You Should Know about the Indonesian Elections," *Pursuit* (blog), February 20, 2019, <https://pursuit.unimelb.edu.au/articles/six-things-you-should-know-about-the-indonesian-elections>.

<sup>12</sup> Nabila Muhamad, "KPU: Pemilihan Pemilu 2024 Didominasi Oleh Kelompok Gen Z Dan Milenial | Databoks," July 5, 2023, <https://databoks.katadata.co.id/datapublish/2023/07/05/kpu-pemilih-pemilu-2024-didominasi-oleh-kelompok-gen-z-dan-milenial>.

<sup>13</sup> Lisye Sri Rahayu, "KPU Sebut Serangan Hoaks Meningkat Pesat di Pemilu 2019," *detiknews*, August 20, 2019, <https://news.detik.com/berita/d-4672954/kpu-sebut-serangan-hoaks-meningkat-pesat-di-pemilu-2019>.

<sup>14</sup> CNBC Indonesia, "Jelang Pemilu 2024, Hoaks Sudah Meningkat 10 Kali Lipat," *CNBC Indonesia*, October 27, 2023, <https://www.cnbcindonesia.com/news/20231027135902-8-484248/jelang-pemilu-2024-hoaks-sudah-meningkat-10-kali-lipat>; Novina Putri Bestari, "Menkominfo Buka-Bukaan Data, Facebook Sarang Hoax Pemilu 2024," *CNBC Indonesia*, October 27, 2023, <https://www.cnbcindonesia.com/tech/20231027114818-37-484199/menkominfo-buka-bukaan-data-facebook-sarang-hoax-pemilu-2024>.

blocked 425 pieces of content containing “hoax” and “hate speech” related to the 2024 election.<sup>15</sup> The government had also blocked online content containing disinformation in the lead up to past elections. For example, Kominfo announced that it had blocked 1,600 social media accounts that spread or contain disinformation one month before the 2019 election.<sup>16</sup>

Actions taken by the government against online disinformation has been carried out on the basis of the Information and Electronic Transactions law (or *Undang Undang Informasi dan Transaksi Elektronik* (UU ITE), here in after “ITE law”). Enacted in 2008, the ITE law was revised in 2016, due to widespread criticism that its overly broad provisions violated the right to free expression.<sup>17</sup> Despite the amendment, however, the ITE law still gave considerable authority to the government to control Internet content.<sup>18</sup> Other regulations that are relevant to tackling disinformation include the Ministry of Communication and Informatics’ Ministerial Regulation No. 5/2020 on Private Electronic System Operators (*Peraturan Menteri Komunikasi dan Informatika Nomor 5 Tahun 2020 tentang Penyelenggara Sistem Elektronik Lingkup Privat*) (hereinafter “Ministerial Regulation (MR) 5”) and the revised Criminal Code (*Kitab Undang-undang Hukum Pidana* (KUHP)). As the ITE law is the law that primarily regulates the Internet, it is also the primary focus of this paper.

This paper begins with an overview of some of the approaches that governments around the world have adopted to regulate disinformation and outline their consequences. Subsequently, existing laws and regulations in Indonesia that have been used to curb the production and dissemination of disinformation are examined and compared against those that have been enacted by other jurisdictions, such as Germany and the European Union (EU). This paper concludes with a discussion on the way forward for Indonesia to better confront the challenges posed by the rapid spread of false information online.

## Different Approaches to Regulating Social Media Companies and Curbing Disinformation

Concerns regarding disinformation have intensified globally since the early 2010s. Events such as the Russo-Ukrainian conflict of 2014, European migrant crisis of 2015, United Kingdom’s (UK) EU membership referendum and United States (US) presidential election in 2016, in conjunction with Germany and France’s elections in 2017, contributed to fears about disinformation, particularly those produced and disseminated by Russia-

---

<sup>15</sup> Banjarmasin Post, “Kominfo Blokir 425 Konten Berkaitan Pemilu, Tribun Network Netral Secara Politik - Banjarmasinpost.Co.Id,” *Tribun News*, November 7, 2023, <https://banjarmasin.tribunnews.com/2023/11/07/kominfo-blokir-425-konten-berkaitan-pemilu-tribun-networ-k-netral-secara-politik>.

<sup>16</sup> Jhon Rico, “Jelang Pemilu 2019, Kemkominfo Blokir 1.600 Akun Penyebar Hoax,” *InfoPublik*, March 17, 2019, <https://infopublik.id/kategori/pemilu/334236/index.html>.

<sup>17</sup> Petra Mahy, Monika Winarnita, and Nicholas Herriman, “Influencing the Influencers: Regulating the Morality of Online Conduct in Indonesia,” *Policy & Internet* 14, no. 3 (2022): 578, <https://doi.org/10.1002/poi3.321>.

<sup>18</sup> Zico Junius Fernando et al., “The Freedom of Expression in Indonesia,” *Cogent Social Sciences* 8, no. 1 (December 31, 2022): 5, <https://doi.org/10.1080/23311886.2022.2103944>.

backed actors, and its impact on democratic processes.<sup>19</sup> In 2018, for example, eighty-three percent of respondents in the Eurobarometer survey stated that “fake news represents a danger to democracy.”<sup>20</sup> Governments’ responses to disinformation concerns have varied, from “launching offensive cyber-operations targeting disinformation actors” to “creating task forces” and drafting new regulations for social media platforms, in addition to criminalizing the creation and dissemination of false information.<sup>21</sup>

Legislative efforts to counter disinformation have increased exponentially,<sup>22</sup> especially during the novel coronavirus 2019 (COVID19) pandemic, when false information contributed to “vaccine avoidance, mask refusal, and [the] utilization of medications with insignificant scientific data.”<sup>23</sup> Consequently, between 2011 and 2022, seventy-eight countries “passed laws designed to limit the spread of false or misleading information on social media.”<sup>24</sup> As laws targeting disinformation are often overly-broad and “allow governments to define prohibited content at their own discretion,” they have been criticized for increasing the risks of infringing on press freedom.<sup>25</sup> Several states, such as Malta, Croatia and Greece, have also established criminal law provisions on disinformation that includes “possible prison sentences upon conviction,” which creates concerns regarding violations on the right to freedom of expression.<sup>26</sup>

Criticism has also been leveled against social media companies for their failures in actively responding to disinformation. For example, a study published by the European Commission in August 2023 found that platforms like Meta (Facebook’s parent company), YouTube, X (formerly Twitter), and TikTok had failed to curb a large-scale Russian disinformation campaign targeting the EU and its allies following the invasion of Ukraine.<sup>27</sup> Previous research by the advocacy group Global Witness also found that Meta had failed to detect and reject advertisements containing hate speech or election disinformation, using

---

<sup>19</sup> Alberto-Horst Neidhardt and Paul Butcher, “Disinformation on Migration: How Lies, Half-Truths, and Mischaracterizations Spread,” *Migrationpolicy.Org* (blog), August 29, 2022, <https://www.migrationpolicy.org/article/disinformation-migration-how-fake-news-spreads>

<sup>20</sup> Eurobarometer, “Fake News and Disinformation Online - March 2018 - - Eurobarometer Survey,” *European Union* (blog), March 2018, <https://europa.eu/eurobarometer/surveys/detail/2183>

<sup>21</sup> Chris Tenove, “Protecting Democracy from Disinformation: Normative Threats and Policy Responses,” *The International Journal of Press/Politics* 25, no. 3 (July 1, 2020): 518, <https://doi.org/10.1177/1940161220918740>; Fernando Nuñez, “Disinformation Legislation and Freedom of Expression,” *UC Irvine Law Review* 10, no. 2 (March 1, 2020): 784.

<sup>22</sup> Roxana Radu, “Fighting the ‘Infodemic’: Legal Responses to COVID-19 Disinformation - Roxana Radu, 2020,” *Social Media + Society* 6, no. 3 (2020), <https://journals-sagepub-com.myaccess.library.utoronto.ca/doi/full/10.1177/2056305120948190>.

<sup>23</sup> Maria Mercedes Ferreira Caceres et al., “The Impact of Misinformation on the COVID-19 Pandemic,” *AIMS Public Health* 9, no. 2 (January 12, 2022): 262–77, <https://doi.org/10.3934/publichealth.2022018>.

<sup>24</sup> Gabrielle Lim and Samantha Bradshaw, “Chilling Legislation: Tracking the Impact of ‘Fake News’ Laws on Press Freedom Internationally,” *Center for International Media Assistance* (blog), July 19, 2023, <https://www.cima.ned.org/publication/chilling-legislation/>.

<sup>25</sup> *Ibid.*

<sup>26</sup> Ronan Ó Fathaigh, Natali Helberger, and Naomi Appelman, “The Perils of Legally Defining Disinformation,” *Internet Policy Review* 10, no. 4 (November 4, 2021), <https://policyreview.info/articles/analysis/perils-legally-defining-disinformation>.

<sup>27</sup> European Commission, “Digital Services Act: Application of the Risk Management Framework to Russian Disinformation Campaigns” (LU: Publications Office of the European Union, 2023), <https://data.europa.eu/doi/10.2759/764631>.

Myanmar, Ethiopia, Kenya, Norway, Brazil, and the US as case studies.<sup>28</sup> Similarly, TikTok's content moderation system was found to be ineffectual, as the platform had approved ninety-percent of advertisements containing election misinformation that were submitted by researchers with Global Witness and New York University.<sup>29</sup>

Social media platforms' efforts on content moderation of disinformation have been hampered by the lack of resources (e.g., human moderators) that they devote to non-English content and users in the Global South, such as in conflict-affected states like Ethiopia and Myanmar,<sup>30</sup> as well as the speed and scale at which disinformation spread. A 2018 study by researchers at the Massachusetts Institute of Technology (MIT) found that "falsehood dilutes significantly farther, faster, deeper, and more broadly than the truth, in all categories of information."<sup>31</sup> Even when social media companies establish partnerships with local fact-checkers and civil society organizations to boost the effort of human moderators, they have been known to ignore requests for support or to incorporate their recommendations, which is an issue that was known with Meta, for example.<sup>32</sup>

Social media companies have pledged to do more to curb disinformation, including through the increased use of artificial intelligence (AI) to flag and take down content before it is reported by users. However, limitations of automated tools continue to exist in understanding the complex "political, social or interpersonal context of speech."<sup>33</sup> Additionally, the *Wall Street Journal* reported in 2021 that Facebook itself was aware that moderation using AI had resulted in only "minimal success" in removing hate speech, violent images, and other problematic content.<sup>34</sup>

Some countries have adopted a voluntary self-regulation model that allows social media companies to govern themselves in terms of how they detect and moderate disinformation. For example, just as the EU had adopted the Code of Practice on

---

<sup>28</sup> Global Witness, "Exposing Social Media Platforms' Failures to Protect Their Users" (Global Witness, December 20, 2022), <https://en.blog/exposing-social-media-platforms-failures-to-protect-their-users/>.

<sup>29</sup> Global Witness, "TikTok and Facebook Fail to Detect Election Disinformation in the US, While YouTube Succeeds" (Global Witness and New York University, October 21, 2022), <https://en/campaigns/digital-threats/tiktok-and-facebook-fail-detect-election-disinformation-us-while-youtu-be-succeeds/>.

<sup>30</sup> Caroline Allen, "Facebook's Content Moderation Failures in Ethiopia," *Council on Foreign Relations* (blog), April 19, 2022, <https://www.cfr.org/blog/facebooks-content-moderation-failures-ethiopia>.

<sup>31</sup> Peter Dizikes, "Study: On Twitter, False News Travels Faster than True Stories," *MIT News | Massachusetts Institute of Technology* (blog), March 8, 2018, <https://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308>.

<sup>32</sup> Jasper Jackson, Mark Townsend, and Lucy Kassa, "Facebook 'Lets Vigilantes in Ethiopia Incite Ethnic Killing,'" *The Observer*, February 20, 2022, sec. Technology, <https://www.theguardian.com/technology/2022/feb/20/facebook-lets-vigilantes-in-ethiopia-incite-ethnic-killing>

<sup>33</sup> Svea Windwehr and Jillian C. York, "Facebook's Most Recent Transparency Report Demonstrates the Pitfalls of Automated Content Moderation," *Electronic Frontier Foundation* (blog), October 8, 2020, <https://www.eff.org/deeplinks/2020/10/facebooks-most-recent-transparency-report-demonstrates-pitfalls-automated-content>.

<sup>34</sup> Deepa Seetharaman, Jeff Horwitz, and Justin Scheck, "Facebook Says AI Will Clean Up the Platform. Its Own Engineers Have Doubts," *Wall Street Journal*, October 17, 2021, sec. Tech, <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184>.



disinformation (CPD) in 2018 (and the follow up 2022 Strengthened Code of Practice on Disinformation) that was signed by tech companies and companies in the advertising industry,<sup>35</sup>

Australia also created the Australian Code of Practice (CoP) on Disinformation and Misinformation in February 2021. The CoP was developed by an industry body known as the Digital Industry Group (DIGI)—whose membership includes Google, Apple, Meta, Twitter and TikTok—which develops a voluntary code of conduct and requires signatories to “identify, assess and address misinformation and disinformation on their services.”<sup>36</sup> Even though the CoP is partially based on the CPD, differences exist in that the CPD is largely focused on disinformation, particularly by state-backed actors, whereas the CoP also addresses content spread by individuals.<sup>37</sup> The CoP, which was updated in December 2022, has been criticized as being inadequate to curb disinformation in Australia. In particular, it lacks strategies on how to tackle disinformation spread by large-scale group messaging (e.g., on WhatsApp) and remains insufficient to engender meaningful transparency from social media companies on how they are combating misinformation and dealing with complaints from the public.<sup>38</sup>

Failures of self-regulation have led to increased efforts by governments to regulate social media companies. Germany, for example, passed the Network Enforcement Act (*Netzwerkdurchsetzungsgesetz* (hereinafter “NetzDG”)) in 2017. Discussions regarding the NetzDG and its eventual passing occurred during a time when anti-refugee hate speech was prevalent online, as a result of the influx of hundreds of thousands of Syrian refugees into Germany in 2015.<sup>39</sup> Simultaneously, multiple states were engaging in efforts to hold social media companies accountable for their content moderation practices (or failures thereof), ranging from creating statutory legislation to establishing ‘co-regulatory’ initiatives and private industry-led associations.<sup>40</sup>

NetzDG mandates the relevant social media companies (i.e., those with more than two million users located in Germany) to remove content that is “clearly illegal” within twenty-four hours after receiving a user complaint, or if illegality is not obvious then the

---

<sup>35</sup> European Commission, “2018 Code of Practice on Disinformation,” June 16, 2022, <https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation>

<sup>36</sup> Australian Communications and Media Authority, “Report on Digital Platforms’ Efforts under the Australian Code of Practice on Disinformation and Misinformation | ACMA” (Australian Communications and Media Authority (ACMA), July 2023),

<sup>37</sup> Edward Hurcombe and James Meese, “Australia’s DIGI Code: What Can We Learn from the EU Experience?,” *Australian Journal of Political Science* 57, no. 3 (July 3, 2022): 297–307, <https://doi.org/10.1080/10361146.2022.2122774>.

<sup>38</sup> Paul Karp, “WhatsApp and Other Messaging Services Left out of New Australian Misinformation Code, Regulator Says,” *The Guardian*, January 8, 2023, sec. Media, <https://www.theguardian.com/australia-news/2023/jan/09/whatsapp-and-other-messaging-services-left-out-of-new-australian-misinformation-code-regulator-says>.

<sup>39</sup> Christoph Hasselbach, “Syrian Refugees Find a Safe Haven in Germany – DW – 03/15/2021,” *Dw.Com*, March 15, 2021, <https://www.dw.com/en/syrian-refugees-find-a-safe-haven-in-germany/a-56872099>.

<sup>40</sup> Robert Gorwa, “Elections, Institutions, and the Regulatory Politics of Platform Governance: The Case of the German NetzDG,” *Telecommunications Policy*, Norm entrepreneurship in Internet Governance, 45, no. 6 (July 1, 2021): 102145, <https://doi.org/10.1016/j.telpol.2021.102145>.

company has seven days to investigate and delete it, or risk fines up to 15 million euros for noncompliance.<sup>41</sup> Crucially, however, NetzDG does not create new categories of illegal content, but enforces twenty-two pre-existing statutes in the German criminal code and requires large social media platforms to apply them to online content.<sup>42</sup> In addition, NetzDG requires social media companies to establish “an elective and transparent procedure” to handle removal complaints and to publish a transparency report on a biannual basis.<sup>43</sup>

The requirement for meaningful transparency in NetzDG was brought about by criticism against social media companies regarding the opaque nature of their content moderation processes. In theory, transparency reports were supposed to provide more insight into their moderation decisions. Yet, research has revealed that none of the transparency reports issued by social media companies “fully adhere to” the Santa Clara Principles, which is a set of principles that aims to help companies achieve meaningful transparency and accountability in content moderation, and that these companies have continued to conceal or obfuscate their internal moderation practices and flagging processes.<sup>44</sup> Furthermore, research has also shown that companies have only become less transparent about their handling of user-generated content and how they enforce their rules or standards.<sup>45</sup> Under the ownership of Elon Musk, for example, X has stopped producing transparency reports that provide details regarding the company’s responses to requests for content restrictions and demands for user data.<sup>46</sup>

NetzDG has been recognized as “the most ambitious attempt by a Western state” to hold social media platforms accountable for moderating online speech, and treated as a template by others seeking to create similar legislation.<sup>47</sup> Since NetzDG’s passing, over twenty countries around the world, including Indonesia, have either adopted or proposed legislation that have “directly or indirectly followed the NetzDG template.”<sup>48</sup> As examined further in the next section, Indonesia’s MR5 is considered to be NetzDG’s “variant,”<sup>49</sup> with

---

<sup>41</sup> Library of Congress, “Germany: Network Enforcement Act Amended to Better Fight Online Hate Speech,” web page, Library of Congress, Washington, D.C. 20540 USA, 2021, <https://www.loc.gov/item/global-legal-monitor/2021-07-06/germany-network-enforcement-act-amended-to-better-fight-online-hate-speech/>.

<sup>42</sup> Heidi Tworek and Paddy Leerssen, “An Analysis of Germany’s NetzDG Law,” Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression (Institute for Information Law, April 15, 2019), <https://dare.uva.nl/search?identifier=3dc07e3e-a988-4f61-bb8c-388d903504a7>.

<sup>43</sup> Google, “Removals under the Network Enforcement Law – Google Transparency Report,” Google Transparency Report, 2023, <https://transparencyreport.google.com/netzdg/youtube?hl=en>.

<sup>44</sup> Aleksandra Urman and Mykola Makhortykh, “How Transparent Are Transparency Reports? Comparative Analysis of Transparency Reporting across Online Platforms,” *Telecommunications Policy* 47, no. 3 (April 1, 2023): 102477, <https://doi.org/10.1016/j.telpol.2022.102477>.

<sup>45</sup> Ranking Digital Rights, “Key Findings from the 2022 RDR Big Tech Scorecard,” *Ranking Digital Rights* (blog), 2022, <https://rankingdigitalrights.org/mini-report/key-findings-2022/>.

<sup>46</sup> Jan Rydzak, “The Stalled Machines of Transparency Reporting – New Digital Dilemmas: Resisting Autocrats, Navigating Geopolitics, Confronting Platforms,” *Carnegie Endowment for International Peace* (blog), November 29, 2023, <https://carnegieendowment.org/2023/11/29/stalled-machines-of-transparency-reporting-pub-91085>.

<sup>47</sup> Tworek and Leerssen, “An Analysis of Germany’s NetzDG Law,” 1.

<sup>48</sup> Jacob Mchangama and Natalie Alkiviadou, “The Digital Berlin Wall: How Germany (Accidentally) Created a Prototype for Global Online Censorship – Act Two” (Copenhagen, Denmark: Justitia, September 2020), <https://globalfreedomofexpression.columbia.edu/publications/the-digital-berlin-wall-how-germany-accidentally-created-a-prototype-for-global-online-censorship-act-two/>

<sup>49</sup> Katitza Rodriguez, “Indonesia’s Proposed Online Intermediary Regulation May Be the Most Repressive Yet,” *Electronic Frontier Foundation* (blog), February 16, 2021, <https://www.eff.org/deeplinks/2021/02/indonesias->

the difference being that Indonesia already has a law that outlines what is considered to be illegal content, which is the ITE law. NetzDG also informed the drafting of the EU's Digital Services Act (DSA), which is a set of rules that came into force in November 2022 and impacts online intermediaries (e.g., access providers), hosts (e.g., hosting services), and platforms (e.g., 'very large online platforms' (VLOP) and 'very large online search engines' (VLOSE), which are those that have more than forty-five million users per month in the EU), as it requires them to assess and mitigate risks, including those emanating from illegal content.<sup>50</sup>

Despite its groundbreaking nature, NetzDG has been criticized for enabling violations to the right to free expression. Firstly, the law mandates companies that host user-generated content to determine when online content violates the law, which is a task that even courts can find challenging, as it requires an understanding of local context, culture, and laws, in addition to case-by-case analysis and investigation.<sup>51</sup> Given NetzDG's tight timelines for content removal and the consequences of fines, the law incentivizes platforms to comply with most complaints, which could lead to the over-removal of content regardless of their actual merits.<sup>52</sup>

And secondly, NetzDG fails to provide judicial oversight, as a court order is not required prior to content takedowns, and lacks a judicial remedy should a social media company's content removal decision violate a user's right to convey or access information.<sup>53</sup> Concerns over NetzDG's overly permissive and broad content removal requirement were validated, as several authoritarian governments quickly followed with their own legislation. For example, two weeks after NetzDG was adopted in Germany, Russia passed its own "copy-paste version" of it, as it explicitly referred to the German law in the legislative process, and this law has since been used in Russia to silence government critics.<sup>53</sup> There are fears that the DSA will result in a similar impact.

Some governments have also pursued 'co-regulatory' approaches to content moderation, which typically combine regulatory frameworks and state oversight with self-regulation by private actors.<sup>54</sup> The DSA, which imposes oversight on platforms' content moderation policies through mandating risk assessments and independent algorithm audits, provides several co-regulatory mechanisms, for example. These mechanisms include trusted flaggers, which are "entities with proven expertise in flagging harmful or illegal content to platforms;"<sup>55</sup> Digital Services Coordinators, who are responsible for the

---

proposed-online-intermediary-regulation-may-be-most-repressive-yet.

<sup>50</sup> Johanna Rinceanu and Randall Stephenson, "Digital Introgenesis," *Eucrim*, no. 1 (2023): 74. <sup>51</sup>

Human Rights Watch, "Germany: Flawed Social Media Law," February 14, 2018, <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>.

<sup>51</sup> Tworek and Leerssen, "An Analysis of Germany's NetzDG Law," 3.

<sup>52</sup> Human Rights Watch, "Germany."

<sup>53</sup> Anu Bradford, "The European Rights-Driven Regulatory Model," in *Digital Empires: The Global Battle to Regulate Technology*, ed. Anu Bradford (Oxford University Press, 2023), 141, <https://doi.org/10.1093/oso/9780197649268.003.0004>.

<sup>54</sup> Amélie Heldt and Stephan Dreyer, "Competent Third Parties and Content Moderation on Platforms: Potentials of Independent Decision-Making Bodies From A Governance Structure Perspective," *Journal of Information Policy* 11 (December 1, 2021): 270, <https://doi.org/10.5325/jinfopoli.11.2021.0266>.

<sup>55</sup> Théophile Hartmann, "EU Digital Services Act: Challenges Remain as Enforcement Begins," *Www.Euractiv.Com*, August 24, 2023, sec. Law enforcement,

DSA's application and enforcement in an EU Member State; and out-of-court dispute settlement bodies," as certified by each Member State's Digital Services Coordinator—all of which are in addition to the threat of fines of up to six percent of the total worldwide annual turnover for VLOPs and VLOSEs.<sup>56</sup>

Although some have praised the DSA for attempting to balance the protection of fundamental rights with the regulation of harmful online content—for example, by protecting users' rights, including "the right to online anonymity and private communication," and by including "some of the most extensive transparency and platform accountability obligations to date"<sup>57</sup>—concerns were still raised regarding its impact on free expression. For example, the DSA allows for national judicial or administrative authorities to issue mandatory takedown orders directly to intermediaries, although this authority is limited to the territory or geographical area of the issuing state, which is a technique known as "geo-blocking."<sup>58</sup> Moreover, there are anxieties regarding political pressures or influence in the issuance of takedown orders, as the European Commission, which oversees the VLOPs and VLOSEs, is an intergovernmental organization, in addition to fears of possible skills shortages to enforce the DSA and conduct independent platform algorithms audits on a regular basis.<sup>59</sup>

Reflecting a similar preference for co-regulation, Australia's Minister for Communications Michelle Rowland released a draft bill in June 2023 that aims to bolster the government's arsenal against harmful online disinformation,<sup>60</sup> particularly by introducing heavy penalties on social media companies. Many organizations and individuals provided their comments to the draft bill during its consultation period, including the Australian Human Rights Commission. In particular, the Commission objected to the use of overly broad and vague terms in the draft bill, such as "misinformation, disinformation, and harm."<sup>61</sup> To avoid violations of fundamental freedoms, the Commission

---

<https://www.euractiv.com/section/law-enforcement/news/eu-digital-services-act-challenges-remain-as-enforcement-begins/>

<sup>56</sup> Josefa Francke, "Can Social Media Councils Tame Digital Platforms?— Digital Society Blog," *HIIG* (blog), September 29, 2022, <https://www.hiig.de/en/social-media-councils/>

<sup>57</sup> Konstantinos Komaitis, Rodriguez Katitza, and Christoph Schmon, "Enforcement Overreach Could Turn Out To Be A Real Problem in the EU's Digital Services Act," *Electronic Frontier Foundation* (blog), February 18, 2022,

<https://www.eff.org/deeplinks/2022/02/enforcement-overreach-could-turn-out-be-real-problem-eus-digital-services-act.>; Asha Allen, "Europe's Big Tech Law Is Approved. Now Comes the Hard Part," *Wired*, accessed December 3, 2023, <https://www.wired.com/story/digital-services-act-regulation/>.

<sup>58</sup> Komaitis, Katitza, and Schmon, "Enforcement Overreach Could Turn Out To Be A Real Problem in the EU's Digital Services Act."

<sup>59</sup> Joan Barata and Jordi Calvet-Bademunt, "The European Commission's Approach to DSA Systemic Risk Is Concerning for Freedom of Expression," *Tech Policy Press* (blog), October 30, 2023, <https://techpolicy.press/the-european-commissions-approach-to-dsa-systemic-risk-is-concerning-for-freedom-of-expression>; Hartmann, "EU Digital Services Act."

<sup>60</sup> Andrea Carson, "More Stick, Less Carrot: Australia's New Approach to Tackling Fake News on Digital Platforms," *The Conversation*, July 14, 2023, <http://theconversation.com/more-stick-less-carrot-australias-new-approach-to-tackling-fake-news-on-digital-platforms-209599>.

<sup>61</sup> Australian Human Rights Commission, "Why Misinformation Bill Risks Freedoms It Aims to Protect," *Australian Human Rights Commission* (blog), August 24, 2023, <https://humanrights.gov.au/about/news/opinions/why-misinformation-bill-risks-freedoms-it-aims-protect>.

argued that “laws targeting misinformation and disinformation require clear and precise definitions.”<sup>62</sup>

Similar to NetzDG, those opposed to the draft bill expressed concern that social media companies would over-censor online content to avoid being fined, which could be up to 1% percent of global turnover.<sup>63</sup> Although the government has delayed the introduction of the legislation into parliament until 2024, if passed, the bill will shift Australia’s social media governance “from a voluntary to a mandatory co-regulatory model.”<sup>64</sup>

An examination of the different approaches to regulating social media companies in their efforts to tackle disinformation has revealed that each approach has its own drawbacks and unintended consequences.

---

<sup>62</sup> Ibid.

<sup>63</sup> Josh Taylor, “Labor to Overhaul Misinformation Bill after Objections over Freedom of Speech,” *The Guardian*, November 13, 2023, sec. Australia news, <https://www.theguardian.com/australia-news/2023/nov/13/labor-misinformation-bill-objections-freedom-of-speech-religious-freedom>.

<sup>64</sup> Ibid.