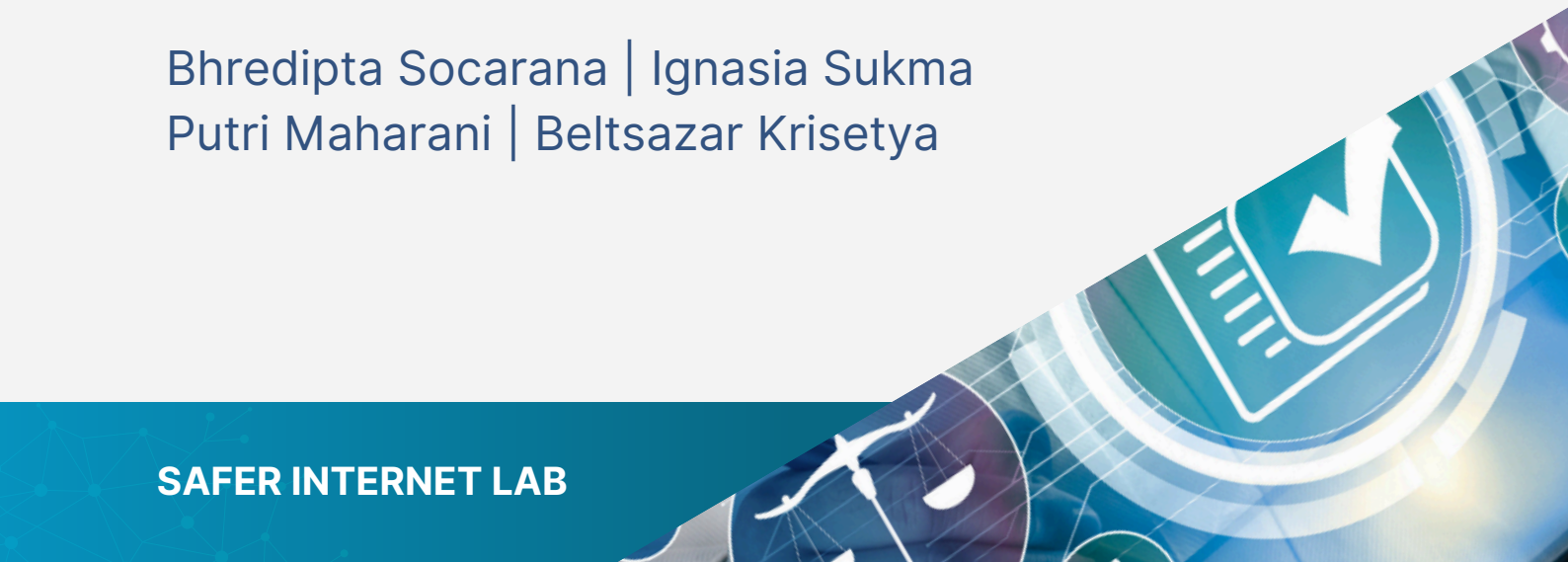
Safer Internet Lab
Led by CSIS in partnership with Google

Centre for Strategic and International Studies

POLICY BRIEF

Advancing Iterative Reform in Indonesia: Regulating "Disturbing Contents" amidst Global Divergence

Bhredipta Socarana | Ignasia Sukma
Putri Maharani | Beltsazar Krisetya



SAFER INTERNET LAB

SAIL Policy Brief a peer-reviewed article series featuring key research findings from Safer Internet Lab researchers and its associates. The views presented are exclusively those of the authors and do not reflect the official stance of SAIL, CSIS, Google, or any other affiliated organisation.

Advancing Iterative Reform in Indonesia: Regulating "Disturbing Contents" amidst Global Divergence

Bhredipta Socarana[1], Ignasia Sukma Putri Maharani[2], Beltsazar Krisetya[3]

Editor: Irene Poetranto[4]

CONTEXT AND PROBLEM STATEMENT

Indonesia's digital governance framework faces a critical juncture as it seeks to reconcile public protection mandates with the imperatives of safeguarding democratic rights and economic growth. In particular, Government Regulation No. 71 of 2019 on Implementation of Electronic System and Transaction (PP 71/2019), enacted to operationalise the 2016 amendment to the Electronic Information and Transactions Law (UU ITE), grants the authorities broad powers to restrict content classified as "disturbing to society" ("meresahkan masyarakat"). However, the regulation's core weakness lies in its use of many ambiguous terms, which conflate legitimate dissent with harmful material.

As the revoking provision to Government Regulation No. 82 of 2012 on Implementation of Electronic System and Transaction (PP 82/2012), PP 71/2019 is a result of an iterative process that evolved through the amendments of UU ITE. PP 71/2019 also provides legal ground for project implementation leading to development of government various content moderation solutions, among others Ministry's Communication and Digital Affair's "Content Moderation Compliance System" (Sistem Kepatuhan Moderasi Konten (SAMAN)).[5]

Albeit the result of an iterative process of more than a half decade, PP 71/2019 remains stipulating provision a nebulous terms, for example term false information and/or facts ("informasi dan/atau fakta yang

[1] Associate Researcher, Safer Internet Lab (SAIL); Expert to Director General of Application Informatics of Ministry of Communications and Informatics Republic of Indonesia (2023-2024)
[2] Project Research Assistant, Safer Internet Lab (SAIL), CSIS
[3] Researcher, Centre for Strategic and International Studies (CSIS); PhD Student, UCL Science, Technology, Engineering, and Public Policy (UCL STEaPP)
[4] Senior Researcher for the Citizen Lab, Munk School of Global Affairs & Public Policy, University of Toronto
[5] The implementation of SAMAN is further stipulated under Minister of Communications and Informatics Decree No. 519 of Implementation Guidelines for Non-Tax State Revenue Originating from the Imposition of Administrative Fine Sanctions for Violations of Obligation Fulfillment by Private Sector Electronic System Operators in the User Generated Content Scope to Perform Access Termination

dipalsukan"), that lacks precise legal parameters. Consequently, it enables subjective interpretations that disproportionately target criticism of public institutions and the targeting of content deemed inconsistent with subjective societal norms. This ambiguity has institutionalised regulatory overreach, such as in the form of redundancy in blocking content that is disturbing to society and pornography.[6]

The lack of clarity in the country's legal framework undermines Indonesia's digital competitiveness, which, despite rising to 43rd globally in the 2024 World Digital Competitiveness Ranking, remains hampered by inconsistent enforcement and investor apprehensions.[7] This is getting more challenging with foreign stakeholders in the digital sector reported unease over the planned expansion of SAMAN, ministry's digital platform to monitor platform's compliance over fine administrative sanction, which imposes penalties without safeguards against erroneous or politically motivated enforcement. This regulatory unpredictability stifles innovation, particularly for platforms hosting user-generated content, as compliance risks in the form of content removal could escalate without procedural transparency. Simultaneously, the conflation of disinformation with criticism erodes public trust, exacerbating tensions between the need to enforce laws to maintain national security and protect civil liberties.

Further complicating this landscape is Indonesia's ambition to become a full member of the Organization for Economic Cooperation and Development (OECD) by 2027, a process that demands alignment with transparency and human rights standards. Current practices under PP 71/2019, however, conflict with these goals, as evidenced by arbitrary restrictions on satirical and artistic expression. The regulation's failure to distinguish between public interest reporting and genuinely harmful content not only jeopardises democratic discourse but also contravenes international benchmarks like the EU Digital Services Act, which emphasises proportionality and judicial oversight. Without reform, Indonesia risks perpetuating a cycle where regulatory ambiguity fuels both economic stagnation and democratic erosion, undermining its "Vision 2045" (VISI Indonesia 2045)

[6] Kementerian Komunikasi dan Digital. (2020, August 26). Tutup konten porno di WhatsApp, polisi surati Kementerian Komunikasi dan Informatika. Kementerian Komunikasi dan Digital. https://www.komdigi.go.id/berita/sorotan-media/detail/tutup-konten-porno-di-whatsapp-polisi-surati-kementerian-komunikasi-dan-informatika

[7] https://www.imd.org/centers/wcc/world-competitiveness-center/rankings/world-digital-competitiveness-ranking/

aspirations of creating an inclusive digital transformation.

Nonetheless, meaningful change in digital governance cannot occur overnight. Reform is inherently an iterative process. Experiences from other countries discussed in this paper demonstrate that there is no silver bullet for achieving a digital ecosystem that is both economically productive and protective of human rights. This paper aims to draw lessons from the diverse approaches adopted by other jurisdictions-such as Singapore's stability-oriented model, the European Union's rights-based Digital Services Act, and the United States' market-driven framework. Rather than positioning Indonesia as inherently right or wrong, the analysis underscores the importance of remaining adaptive and iterative. By examining how different countries reconcile competing priorities-public order, and freedom of expression, this study seeks to inform Indonesia's ongoing regulatory evolution. The objective is to encourage a process of continuous learning and recalibration, ensuring that Indonesia's framework remains responsive to both domestic needs and shifting global standards.

ANALYSIS OF REGULATORY CHALLENGES

Domestic Issues

Indonesia's current regulatory framework for addressing "disturbing content" under PP 71/2019 suffers from systemic flaws rooted in its ambiguous definitions and enforcement practices. This vagueness not only undermines freedom of expression but also fosters a culture of regulatory overreach, where content moderation to take down requests submitted by Government authorities is often assumed to prioritise reputational protection over public safety. The planned expansion of the SAMAN platform by adding more content categories including erroneous or politically motivated content to be monitored with SAMAN, which through SAMAN government will monitor platforms' compliance to government 1 x24 hours takedown requests otherwise at the risk of being sanctioned, exacerbates these risks. Without procedural transparency or appeals mechanisms, platforms face heightened pressure to comply with takedown orders, eroding trust in Indonesia's digital governance framework.

Further compounding these challenges is the inconsistent application of content restrictions across ministries. While gambling-related contents are available to be taken-down under the legal basis of online gambling prohibition content, it is

also frequently flagged by Komdigi under PP 71/2019 as a public disturbing content.[8] It creates enforcement activities to be varied based on subjective interpretations of institutions submitting the takedown request to the Ministry rather than objective legal criteria. For instance, the use of "content disturbing to society" category is used by the Ministry of Women Empowerment and Child Protection[9] to take down content-related to LGBTQ+ (lesbian, gay, bisexual, transgender, queer/questioning) materials. This inconsistency deters legal certainty which affect foreign investment, as businesses face unpredictable compliance burdens as it might lead to arbitrary restrictions on user-generated content. The absence of clear guidelines for cross-ministerial coordination further fragments enforcement, creating loopholes for misuse while hindering cohesive policy implementation.

It is getting more challenging with the lack of clarity on the room for the government to respond to the dynamic of content development at present. Whilst the PP 71/2019 has attempted to accommodate the scope of content under many of its provisions (i.e., disturbing content provisions), several provisions remain wide-spectrum and too vague, leaving uncertainty for society on how to be compliant while enjoying their fundamental rights. With the upcoming implementation of the regulatory framework such as Law No. 1 of 2023 on Criminal Code prohibiting defamatory statements toward state's institution.[10] As such, having a more iterative governance approach on content moderation becomes even more pertinent.

COUNTRIES' BENCHMARKS

Comparative analysis of global frameworks reveals critical gaps in Indonesia's approach. The EU Digital Services Act (DSA), for instance, distinguishes between illegal content (e.g., terrorism, child exploitation) and harmful content (e.g., disinformation), as well as mandating the publication of transparency reports, mechanisms for user appeals, and independent oversight. In contrast, PP 71/2019's broad categorisation of "disturbing content" fails to differentiate satire, artistic expression, or public interest reporting from genuinely harmful material. Singapore's Protection from Online Falsehoods and Manipulation Act (POFMA)

[8] https://www.bloombergtechnoz.com/detail-news/70175/komdigi-blokir-1-3-juta-situs-judi-online-tujuh-bulan-terakhir/2

[9] Kementerian Pemberdayaan Perempuan dan Perlindungan Anak (KemenPPPA). (2023, August 24). Cuplikan video lagu anak diduga berunsur LGBT, KemenPPPA ambil langkah. KemenPPPA. https://www.kemenpppa.go.id/page/view/NDY3Nw==#

[10] https://icjr.or.id/revisi-uu-ite-2024-dan-kuhp-2023-tentang-berita-bohong-penghinaan-dan-ujaran-kebencian-harus-dilakukan-merespons-berbagai-putusan-mk-tentang-kebebasan-berekspresi/

offers another model, defining actionable falsehoods as verifiably inaccurate claims causing demonstrable public harm. Its corrective system—provides a less overreach and oversimplification approach toward falsehoods, by enabling the Singaporean authority a range of action from requiring platforms to label disputed content to platform blocking —attempts to balance public protection with free speech - which quite different with Indonesia's approach such as takedown request, imposition of administrative sanctions, and ordering ISP to block access to the internet platform in case of platform's non-compliance.

Malaysia's draft Online Safety Bill introduces a risk-based classification system, obliging platforms to implement tailored mitigation strategies (e.g., age verification, content labelling) instead of blanket takedowns. At present, Malaysia adheres to industry guidelines which was enacted based on consultation with industry under the relevant ministry. Meanwhile, Brazil's Civil Rights Framework for the Internet (*Marco Civil da Internet)* emphasises judicial oversight for content restrictions, ensuring due process—a safeguard absent under PP 71/2019 -- and POFMA as well as Malaysia's draft Online Safety Bill, where ministerial directives often bypass judicial review.

The EU and ASEAN's emphasis on multi-stakeholder governance further highlights Indonesia's institutional gaps. The DSA requires platforms to collaborate with civil society and academia in risk assessments, which fosters democratic accountability.

POLICY RECOMMENDATIONS

1. Legal and Definitional Reforms

The cornerstone of Indonesia's regulatory overhaul lies in dismantling the ambiguity of PP 71/2019's provision regarding "content disturbing to society" ("konten meresahkan masyarakat") by adopting a bifurcated framework that distinguishes between illegal and harmful content.[11] Current provisions conflate legitimate dissent, public interest reporting, and subjective societal norms with genuinely dangerous material, creating a governance vacuum ripe for misuse. One of our findings indicated the use of the category of "content disturbing to society" to target LGBTQ+ (lesbian, gay, bisexual, transgender, queer/questioning) content by one of Indonesian Ministries albeit the absence of empirical proof that such content resulted in disturbance to society.[12]

[11] The distinction between illegal and harmful content is largely based on the 2024 EIT Law revision which adds harmful content distribution as a criminal offense.

[12] Kementerian Pemberdayaan Perempuan dan Perlindungan Anak (KemenPPPA). (2023, August 24). *Cuplikan video lagu anak diduga berunsur LGBT, KemenPPPA ambil langkah*. KemenPPPA.

To align with global practice, Indonesia must legislatively redefine restricted content into two precise categories. Illegal Content should encompass material explicitly prohibited under criminal law, such as incitements to violence, child exploitation, or terrorism-related propaganda. This category requires no interpretive flexibility, ensuring enforcement aligns with universally recognised crimes. Exemptions for satire, artistic expression, and public interest journalism must be codified to prevent the weaponization of PP 71/2019 against democratic discourse. The EU's carve-outs for parody and Singapore's corrective-notice model for disputed content offer references. Legislative reform should further mandate procedural transparency, judicial oversight, requiring the Ministry of Communication and Digital Affairs ("Komdigi") to publish granular guidelines on categorising content, periodically publish contents they have taken down and establishing a multi-stakeholder committee—comprising legal experts, civil society, and technologists—to issue public recommendation based on their audit towards Komdigi's enforcement patterns and mitigate bias.

Integrating these definitions into Komdigi's operation will prevent Komdigi from imposing administrative sanction fees (and later access blockage) against content which restricts freedom of expression. This on the other hand would help to distribute risk from Komdigi to other stakeholders (i.e., civil society, and social media platforms) where every stakeholder is responsible for content moderation efforts. . By anchoring reforms in such specificity Komdigi can transform PP 71/2019 from a tool of reputational shielding into a framework fostering both digital safety and democratic vitality.

2. Procedural Safeguards

Indonesia's current content moderation framework lacks mechanisms to ensure due process, transparency, and accountability, creating systemic risks of overreach and public distrust. Establishing a tiered procedural system to balance swift enforcement with fundamental rights protections could provide a better oversight for the society over government's action.

Drawing from the EU Digital Services Act's transparency mandates and Singapore's corrective notice model, Indonesia should implement a three-stage moderation protocol, beginning with notification and correction. Platforms would be required to notify creators of flagged content within 24 hours, provide specific grounds for the complaint, and allow a 48-hour window for

https://www.kemenpppa.go.id/page/view/NDY3Nw==#

response or voluntary removal. The transparency mandates also extend to mandate platforms being transparent for a certain period of time on the removal effort they carried out. This way the public can have a better understanding on how the removal activities took place, which parties lodged the complaint, under what basis, and other information to ensure content removal is done legitimately. Considering the public interest of this mandate, the regulator must ensure any revision to PP 71/2019 protects the transparency mandate carried out by platform as stipulated under Santa Clara Principle.

To address power imbalances, an Independent Review Body comprising civil society representatives, legal experts, and digital rights advocates should be established under revision of PP 71/2019 as a redress-seeking mechanism to adjudicate disputed takedowns. This body would evaluate whether contested content meets the revised definitions of "illegal" or "harmful" material, preventing subjective interpretations by state actors. For instance, advocacy content flagged under vague "public order" grounds could be reassessed against objective harm criteria. The body shall not replace the court's authority to determine the legality of content removal. Instead, this body's decision could serve as interpretation of the legal provision to the variety of content and its legality to be removed from the internet which act as guidelines for Komdigi and stakeholders (especially platform) in moderating content online. The review process should mirror Brazil's judicial oversight model under Marco Civil da Internet, requiring state agencies to obtain court approval including for politically sensitive takedowns.

Transparency reporting must become part of the revised substantive part of PP 71/2019, obliging platforms to publish quarterly metrics on government requests, compliance rates, and appeals outcomes, adhered by Komdigi and Government Agency, and overseen by the public which enforcement mechanism involves public information inquiry under Law 14 of 2008 on Public Information Openness, and applicable administrative law proceeding. These reports should detail the ministries initiating takedowns, content categories involved, and resolutions—mirroring the EU DSA's standardised reporting framework.

3. Institutional Coordination

Komdigi must centralize oversight by developing standardized guidelines for cross-ministerial collaboration, ensuring alignment between content removal directives and narrow legal criteria, and legally mandating all content removal requests to be made through its centralized platform. This requires establishing a Multi-Stakeholder Oversight Committee comprising civil society representatives,

legal experts, and digital rights advocates to audit enforcement patterns and mitigate bias. The committee would operate similarly to Brazil's judicial review model under *Marco Civil da Internet*, mandating inter-ministerial consultations for contested takedowns and preventing unilateral actions that conflate public criticism with threats to societal stability (See procedure part). Additionally, Komdigi could raise an argument, failure of cross-ministerial collaboration to follow Komdigi's centralised removal request may lead to state's loss of potential revenue, considering content removal violation by platform and electronic system is considered as non-tax state revenue. Both requirements and arguments must be included as part of the PP 71/2019 revision to provide a legal binding effect.

To harmonize enforcement, capacity-building programs for law enforcement and the judiciary are critical, although the larger goal is for every stakeholder involved in policy-making and policy implementation. For instance, training modules on digital rights principles could address systemic misinterpretations of vague provisions. These programs should emphasise distinguishing legitimate dissent from genuinely harmful content, drawing on case studies from the EU Digital Services Act's (DSA) transparency protocols. Additionally, Komdigi should mandate in its PP 71/2019 revision, periodical review to all regulations related to content moderation from PP 71/2019 and its implementing regulation both in substantive and procedural manner involving meaningful participation of members of the public including court authority, civil society, industry, and academics.

Finally, aligning these reforms with Indonesia's OECD accession goals requires adopting participatory governance frameworks to synchronize regulatory priorities with international standards on transparency and human rights. By institutionalising collaboration through inserting legal and definitional reforms, procedural safeguards, and institutional coordination as part of revising PP 71/2019, Indonesia can transform PP 71/2019 from a tool of reputational protection for Indonesia as a nation not just to protect government's reputation into a mechanism which fosters both society's digital safety and democratic accountability.

IMPLEMENTING CHANGES

The transformation of Indonesia's content governance framework demands a phased approach that balances urgency with institutional readiness through having incorporated revisions under PP 71/2019 following the momentum of UU ITE revision. Drawing from the 2021–2024 Digital Indonesia Roadmap and lessons from the

SAMAN platform's pilot phase, reforms must prioritise stakeholder collaboration, technological integration, and iterative evaluation. Effort must be prioritised not just to create a new legal instrument, but to foster the effectiveness of the existing laws by way of enabling society to have a clearer legal provision on content moderation (i.e., replacing public disturbance content provision into a more specific provision), a more transparent system to oversee content removal effort done by government agencies and platform, as well as a better redress mechanism.

Short to Medium Term Priorities (0–18 Months)

Immediate action should focus on removing "disturbing content" provision under PP 71/2019. To rectify this, Komdigi should convene multi-stakeholder workshops involving civil society, legal experts, and digital platforms to codify two distinct categories: illegal content (explicitly prohibited under criminal law, e.g., incitement to violence, child exploitation) and harmful content (material requiring graduated responses, e.g., public health disinformation). This bifurcation aligns with OECD harm taxonomies while preserving exemptions for satire and public interest journalism, as seen in the EU DSA's carve-outs

Concurrently, Indonesia must pilot a tiered moderation system for "very large platforms" (VLOPs) under thresholds mirroring the DSA's user-base criteria. Platforms like TikTok and Facebook would be required to notify creators of flagged content within 24 hours and allow 48-hour rebuttal windows, a mechanism tested in Singapore's POFMA corrective notice system.

Medium to Long Term Priorities (6–18 Months)

Building on initial pilots, institutionalize the Independent Review Body comprising judicial experts, human rights advocates, and information integrity specialists to adjudicate escalated cases. Furthermore, Indonesia's regulatory framework must evolve from punitive enforcement to fostering collaborative stewardship. This can be done by providing incentives to the platform through guaranteeing certain immunities toward its transparency mandate, award for compliance, or legalised (and institutionalised) policies encouraging society and every stakeholder over the internet to ensure checks and balances of each role on the internet. This could serve not just to ensure fundamental human rights protection but also risk mitigation. In doing that, it is imperative as part of the collaborative stewardship development policymaking and policy implementation capacity building to be implemented for stakeholders involved in the policy-related lifecycle. This way all

relevant policies from end-to-end perspective could be mitigated by being restrictive toward freedom of expression.

CONCLUSION

Indonesia's efforts to regulate "disturbing content" under PP 71/2019 must evolve to address systemic risks to democratic discourse, digital rights, and economic growth. The regulation's current vagueness—exemplified by subjective enforcement against criticism and satirical expression—undermines both public trust and investor confidence.

Reforms centred on precise definitions, procedural fairness, and technological accountability are not just necessary but urgent.

These steps, coupled with Indonesia's iterative reform path offers a third way between the EU's procedural rigidity and Singapore's efficiency focus. By treating PP 71/2019 as a living framework—subject to citizen's led audit and feedback on the implementation of PP 71/2019 based on a transparency report published by platform—Indonesia can reconcile global norms with local realities. Citizen-led audits under Law 14/2008 on Public Information Openness would crowdsource regulatory impact analyses, while cross-ministerial labs pre-empt enforcement fragmentation. This approach positions Indonesia not as a passive norm-taker but as a pioneer of adaptive digital governance, crucial for navigating the geopolitical stakes of content regulation in the 21st century.

ANNEXES

Annex 1: Comparative Table

Elements	Indonesia	European Union (EU)	Singapore
	Government Regulation No. 71 of 2019 (PP 71/2019)	Digital Service Act (DSA)	Protection from Online Falsehood and Manipulation Act (POFMA)
Definition of restricted contents	Vague: “Informasi dan/atau fakta yang dipalsukan” (false information/facts) with no clear parameters.	Not defined in the regulation but instead provides the definition of “illegal content”.	Communication of false statement of fact in Singapore.
False information	Described vaguely as “informasi dan/atau fakta yang dipalsukan” (false information) without clear legal parameters.	The 2022 Strengthened EU Code of Practice on Disinformation acknowledges the role of ‘misleading or outright false information’ as an integral part of disinformation. However, it does not separately define ‘false information’ as an independent legal or technical category within the framework of the Code.	- Section 2(2)(b) defines a false statement as "if it is false or misleading, whether wholly or in part, and whether on its own or in the context in which it appears". - Particularly, a person is prohibited from committing any acts or creating a statement, which that person knows or has the reason to believe that it is a falsehood. This applies both, within or outside of Singapore, and it falls under Section 7 of POFMA.
Legal basis for restrictions	Broad references to “public order” and “societal norms,” under Law Number 1 of 2024 on Electronic Information and Transactions (ITE Law)[13] and	Social media platforms are required to moderate content under EU Law (The DSA), while simultaneously upholding	Relevant Ministers may issue Directives under Part 3 and Part 4 of the POFMA. Additionally, Section 48 also empowers Authorities to issue a Code of Practice to support enforcement.

[13] Article 40A (5) of Law Number 1 of 2024 on Electronic Information and Transactions (ITE Law).

	Government Regulation No. 71 of 2019 on the Implementation of Electronic Systems and Transactions[14], enabling subjective enforcement.	ECHR safeguards.[15]	
Content categories	Ambiguous conflation of criticism, disinformation, and moral violations.	Illegal content and systemic risks are key concerns, especially for Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs). The 2018 Code of Practice on Disinformation, along with its 2022 strengthened version, outlines practical measures to address and mitigate the spread of disinformation.	- Statement of facts, which a reasonable person is seeing, hearing, or perceived as representation of facts. - POFMA is not intended to cover opinions, criticism, satire, or parody, as it targets falsehoods, not opinions or criticisms.[16]
Procedural safeguards	No mandated appeals process.	Out of court dispute resolution to certified bodies. The European Commission publishes and updates a list of certified out-of-court dispute settlement bodies.	An appeal against a POFMA Direction or Order must first be submitted to the Minister who issued the directive. If the Minister rejects the request, the next step is to file an appeal with the General Division of the High Court.
Transparency requirements	Limited public reporting.	Providers of online platforms are required to submit a statement of reasons for content moderation decisions to the DSA Transparency Database[17]	POFMA Code of Practice for Transparency of Online Political Advertisement outlines transparency obligations for digital advertising intermediaries and internet intermediaries.[18]
Exemptions	No explicit protections for satire, art, or	The DSA applies broadly to all	Not applicable to opinions, debates, legitimate

[14] Article 96 (b) of Government Regulation No. 71 of 2019 on the Implementation of Electronic Systems and Transactions.
[15] Article 8 *juncto* Article 10(2) of European Convention on Human Rights (ECHR).
[16] *The Online Citizen Pte Ltd v Attorney-General and Another Appeal and Other Matters,* [2021] SGCA 96 (Sing.).
[17] https://transparency.dsa.ec.europa.eu/
[18] https://www.pofmaoffice.gov.sg/files/documents/political%20advertisements%20code%20and%20annex.pdf

	public interest reporting.	intermediary services that operate in the EU.	discourse, or content of a satirical, humorous, or subjective nature.
Stakeholder involvement	Centralized enforcement; limited public and civil society input.	Enforced by The EU Commission and Digital Service Coordinators.	Vested in the respective Ministers or an Alternate Authority designated by the Minister.[19]

[19] Section 6 *juncto* Section 53(1) of the POFMA.

Safer Internet Lab

saferinternetlab.org

Jl. Tanah Abang III no 23-27
Gambir, Jakarta Pusat. 10160

Find Us On

CSIS Indonesia | Safer Internet Lab