

Audrine, Pingkan; Setiawan, Indra

Research Report

Impact of Indonesia's Content Moderation Regulation on Freedom of Expression

Policy Paper, No. 38

Provided in Cooperation with:

Center for Indonesian Policy Studies (CIPS), Jakarta

Suggested Citation: Audrine, Pingkan; Setiawan, Indra (2021) : Impact of Indonesia's Content Moderation Regulation on Freedom of Expression, Policy Paper, No. 38, Center for Indonesian Policy Studies (CIPS), Jakarta

This Version is available at:

<https://hdl.handle.net/10419/249418>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

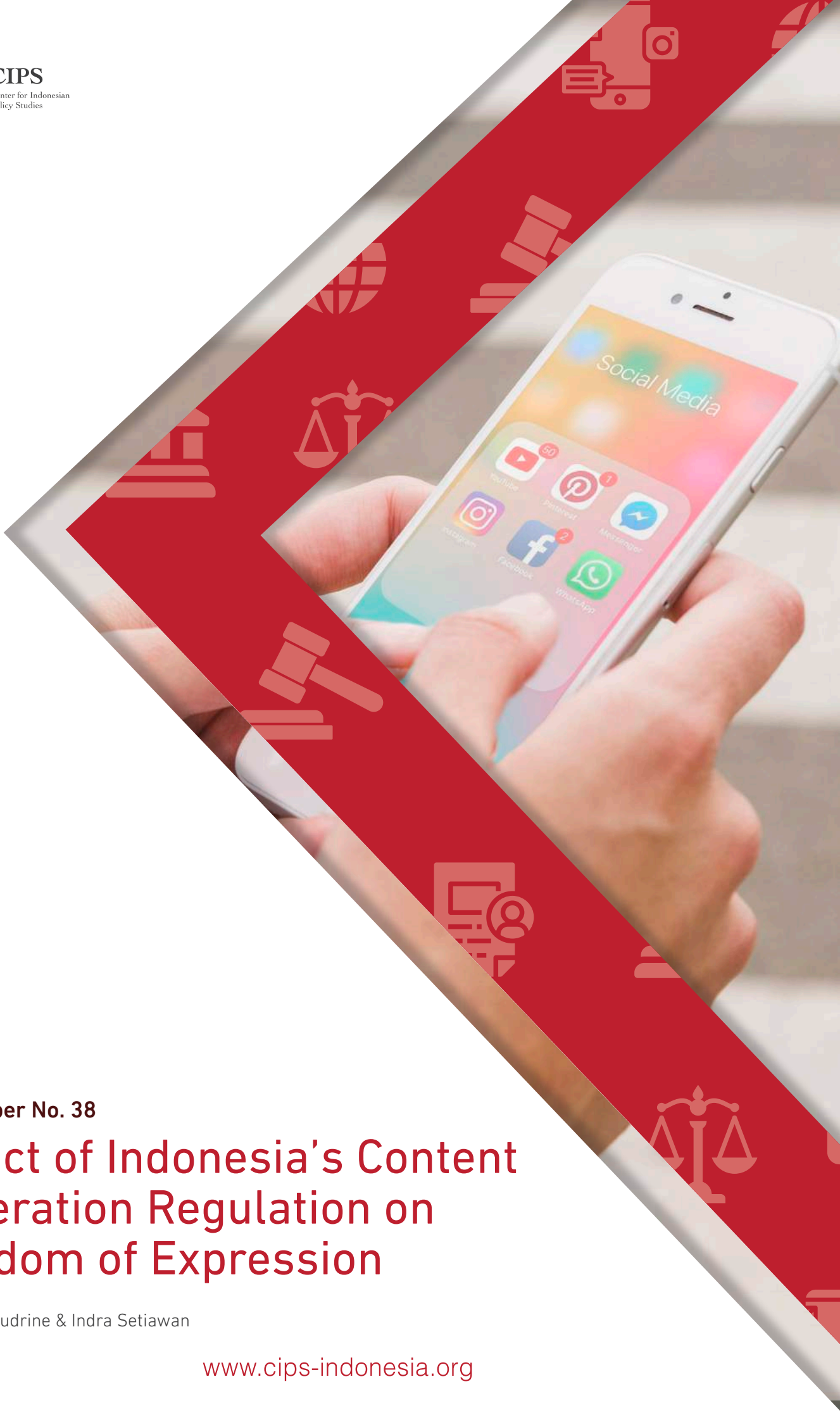
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



CIPS
Center for Indonesian
Policy Studies



Policy Paper No. 38

Impact of Indonesia's Content Moderation Regulation on Freedom of Expression

by Pingkan Audrine & Indra Setiawan

www.cips-indonesia.org



Policy Paper No. 38
**Impact of Indonesia's Content Moderation Regulation
on Freedom of Expression**

Authors:
Pingkan Audrine & Indra Setiawan
Center for Indonesian Policy Studies (CIPS)

Jakarta, Indonesia
July, 2021

Acknowledgment

We also thank Thomas Dewaranu for his assistance in this research

Copy Editor: Janet Bufton

Cover: [freepik.com/freepik](https://www.freepik.com/freepik)

CONTENT

Executive Summary	6
Content Moderation Landscape	7
The Need for Content Moderation.....	7
Content Moderation and Freedom of Speech: A Global Debate.....	8
Content Moderation Debate in Indonesia.....	9
The Impact of the Lack of Due Process in MOCI Regulation	
No.5/2020 on Users’ Freedom of Expression	10
Content Moderation Under MOCI Regulation No. 5/2020.....	10
The Concentration of Power in the Government and the Lack of Due Process.....	10
Challenges in Defining “Prohibited Content” Underscores the Need for Due Process.....	11
Increasing Global Demand for Due Process in Content Moderation.....	12
Limited Content Moderation Action.....	13
The Impact of the Lack of Due Process in MOCI Regulation No. 5/2020 on Content Moderation Mechanism for ESOs	14
ESOs’ Dilemma in Responding to MOCI’s Takedown Requests Within a Restricted Time Span.....	14
User Generated Content Platform Liability Under Self-initiated Content Moderation.....	15
Strengthening the Safe Harbor for ESOs to Strengthen Public-Private Collaboration.....	15
Recommendations	16
References	18

EXECUTIVE SUMMARY

Indonesia is set to implement regulation aimed at improving online content moderation that is potentially problematic for both users and private electronic system organizers (ESOs) through the Ministry of Communications and Informatics (MOCI) Regulation No. 5/2020 on Electronic System Organizers in the Private Sector. This regulation requires that ESOs remove or block content at MOCI's behest within a strict time span—24 hours for non-urgent content and four hours for urgent content.

The lack of due process for ESOs trying to meet their obligations when they receive takedown requests from MOCI is exacerbated by a loose definition of “prohibited content.” The regulation creates a risk of unaccountable over-enforcement that violates individuals' freedom of expression in the digital space.

While websites need content moderation, the timeframes set by MOCI make it impossible for private ESOs to properly assess takedown requests before carrying them out. This creates a dilemma for private ESOs—they can quickly and without proper evaluation take down the content, in which case they risk being accused of violating individual rights, or they can refuse the request, which could lead to access blocking by the MOCI.

These problems illustrate why MOCI must adopt a co-regulation approach and involve private ESOs in the development of implementation guidelines. These guidelines should include due process for content removal orders by the government that includes an appeal mechanism, a reasonable timeframe for compliance, and immunity rights from damage that may arise from users' posts in platforms for ESOs that have met their obligations under MOCI Regulation No. 5/2020.

A crucial missing piece from MOCI Regulation No. 5/2020 is a procedure for platforms to object to content removal requests. The establishment of an independent oversight board chaired by representatives from the government, private ESOs, and civil society organizations to resolve disputes over content moderation should be considered to ensure that moderation is conducted while balancing concerns about freedom of speech.

MOCI should refrain from defining content moderation merely as ‘content removal’ from online platforms and instead allow for a broader definition of the term. Rather than leaving private ESOs with two extreme choices—remove or keep—over the allegedly prohibited content, MOCI Regulation No. 5/2020 needs to acknowledge that actions such as demoting, downranking, demonetizing, content labelling, limiting access, and warnings for users about possibly sensitive and harmful posts, are already-existing mechanisms that help keep the internet safe and usable. These mechanisms should be included as content moderation actions by the regulation.

Adopting a co-regulation approach to develop a due process for content moderation and allowing a broader set of actions by ESOs in response to removal requests would facilitate a more balanced approach to internet safety and freedom of speech.

CONTENT MODERATION LANDSCAPE

Internet use is growing quickly everywhere in the world and Indonesia is no exception. Almost 50% of the Indonesian population are internet users who use digital platforms ranging from social media, digital marketplaces, knowledge sharing, and entertainment offered by electronic systems providers (Statistics Indonesia, 2020). Indonesian laws, including Ministry of Communications and Informatics (MOCI) Regulation No. 5/2020—the focus of this policy paper, refer to these providers as electronic system organizers (ESOs). MOCI Regulation No. 5/2020 defines ESOs as any person, state or business entity, or community that provides, manages, and/or operates electronic systems individually or collectively for their users, for their interests, or the interest of other parties. This definition is all-encompassing and includes all sizes of private online platforms operating domestically or from overseas.

Information exchange between users is continuous and fast-paced, especially on user-generated content (UGC) platforms or intermediaries that rely on electronic contents created by their users in their platforms. MOCI Regulation No. 5/2020 describes that these contents include but are not limited to writing, sound, images, maps, designs, photos, electronic data interchange, electronic mail, telegram, telex, telecopy, letters, access codes, symbols.

Social media is one example of a UGC platform with a vast and continuous digital information flow of UGC. Every minute, there were 41 million WhatsApp messages shared, 300 thousand Instagram stories uploaded, and 300 thousand messages and images posted on Facebook worldwide in 2020 (Statista, 2020).

Statistics Indonesia's (2019) data show that social media is a reason to access the internet for 87% of Indonesian internet users. As of January 2021, 170 million Indonesians, roughly half of the country's total population, are social media users (Statistics Indonesia, 2021; We Are Social & Hootsuite, 2021). Among Indonesian adults, YouTube, WhatsApp, Instagram, Facebook, and Twitter are the most popular, used by 94%, 88%, 87%, 86%, and 64% of adult internet users in Indonesia, respectively (We Are Social & Hootsuite, 2021).

Statistics Indonesia's (2019) data show that social media is a reason to access the internet for 87% of Indonesian internet users.

The Need for Content Moderation

The type of content posted varies greatly and carries different consequences. Along with benefits from content that can be educational, informative, enriching and free, internet users can also find pornographic, hateful, violent, abusive, and obscene content that can have serious personal and social costs.

The risk is even bigger for UGC platforms, where users' posts are at the heart of their business model. While they may help create "networked publics" (Gillespie, 2018), they may also be prone to online harassment, bullying, revenge pornography, and other damaging misuse of their platforms. A mechanism is needed to decide what stays online and what gets taken down. This is known as content moderation (Barret, 2020). Many social media platforms practice self-regulation by developing their own community standards and censoring posts that fail to meet them.

While weak content moderation comes with a risk of harmful material, excessive content moderation may turn into excessive censorship that could violate individual freedom of expression. Over-enforcement of content moderation could balloon into restriction of public opinion. MOCI Regulation No. 5/2020 attempts to regulate content moderation in a way that falls between these extremes, but its problematic enforcement mechanism means that it errs on the side of violating freedom of expression.¹

The enforcement of MOCI Regulation No. 5/2020 has been delayed by the passage of MOCI Regulation No. 10/2021. This paper aims to address some of the issues from both the private ESOs and the users and suggest improvements to the regulation before it becomes fully enforceable.

Content Moderation and Freedom of Speech: A Global Debate

The discourse of liberation versus restriction and where to draw a line between them in internet censorship is the subject of a global debate. This debate surrounds questions of how content moderation should balance protecting freedom of speech and maintaining safe online activity and of what roles digital platforms should play.

In many countries, there is growing demand from internet users and civil society organizations for improved protection for users from harmful online content.

In many countries, there is growing demand from internet users and civil society organizations for improved protection for users from harmful online content (Vogels, 2021; Lomba et al., 2021). However, to what extent platforms should be held responsible for the material on their sites varies widely across different jurisdictions.

Section 230 of the United States Communication Decency Act, for example, guarantees immunity for platforms from liability for any published information or content by third-party users. While some welcome this for the freedom that it allows internet users who are developing content and applications, others worry that the network effects, especially in social media, accelerate the spread of harmful content, such as misinformation (Kerry, 2021).

The polar opposite is the onerous content restrictions in China. China is well-known for banning both domestic and foreign digital platforms that fail to comply with its censorship requirements, including blocking politically sensitive content. Even a tech giant like Apple has been forced to remove up to 55,000 applications from the App Store since 2017 in order to access Chinese consumers (Caster, 2021)

¹ It should be noted that there are other concerns over MOCI Regulation No. 5/2020, particularly regarding data governance, but this policy paper exclusively looks at its content moderation provisions and how the lack of due process therein may be damaging to the users and the private ESOs in Indonesia.

Germany's Network Enforcement Act (NetzDG), though not as restrictive as China, also requires social networks with more than 2 million users to take down illegal content within a restricted time and imposes heavy fines for noncompliance (De Streef, et al., 2020). A recent reform in the federal parliament is reported to have added a responsibility for social networks to file reports to the federal police for criminal content on their platform (Deutscher Bundestag, 2020).

According to SAFEnet (2021 a), Indonesia's content moderation under MOCI Regulation No. 5/2020 reflects the approach of NetzDG, especially in the strict time span to comply with takedown requests and the strong concentration of power in the MOCI.

Content Moderation Debate in Indonesia

Though content moderation enforcement has only recently gained special attention from platforms and the public in Indonesia, debates about internet censorship have been ongoing for at least a decade. For example, many expressed concern over the use of Law No. 11/2008 on Electronic Information and Transaction and its 2016 revision (EIT Law) to repress political views and beliefs. Access termination and internet shutdown by the government on the grounds of preventing misinformation were often met with pushbacks from the public. The judiciary branch has even ruled unlawful the controversial internet access shutdown in Papua by the government in 2019 (Adjie, 2020).

Rather than reconsidering state involvement in the digital space, MOCI Regulation No. 5/2020 reemphasizes MOCI's central role in Indonesia's digital governance by establishing MOCI's authority to request that private ESOs take down prohibited content with limited due process and no appeal mechanism.

MOCI Regulation No. 5/2020 mandates that all ESOs comply with takedown requests by MOCI within a restricted time period—24 hours for non-urgent prohibited content, and four hours for urgent prohibited content. This makes it difficult for the ESOs to comprehensively check the reported content, especially when the regulation gives no room for the private ESOs to express disagreement with the requests. For UGC platforms, this obligation is in addition to operating their own content moderation mechanisms.

Access termination and internet shutdown by the government on the grounds of preventing misinformation were often met with pushbacks from the public.

THE IMPACT OF THE LACK OF DUE PROCESS IN MOCI REGULATION NO.5/2020 ON USERS' FREEDOM OF EXPRESSION

Content Moderation Under MOCI Regulation No. 5/2020

MOCI Regulation No. 5/2020 is an implementing regulation derived from Government Regulation (GR) No. 71/2019 on the Implementation of Electronic Systems and Transaction, which specifically governs private ESOs. MOCI Regulation No. 5/2020 mandates all types of private ESOs, regardless of their size, to keep their platforms free from prohibited content—defined as content that violates existing legislation, content that disturbs public order, or content that provides access to prohibited content. The public, courts, legal officers, and government agencies can file reports of prohibited content to MOCI. MOCI then verifies whether the content is prohibited and passes on takedown requests to the ESOs. Upon receiving the request from MOCI, an ESO has up to 24 hours to take down content deemed non-urgent, or four hours for content deemed urgent. Urgent content includes materials related to terrorism, child pornography, and content that creates a public disturbance or disrupts public order.

In addition to this, operators of UGC platforms must develop their own reporting mechanism whereby the public can actively participate in spotting and reporting prohibited content. Upon receiving a report, the ESOs must verify them and decide whether the materials are violating the standards set by the MOCI Regulation No. 5/2020 as well as the platform's community guidelines.

The Concentration of Power in the Government and the Lack of Due Process

In the broadest sense, due process can be understood as a procedural protection to ensure fairness and protection of private rights in a particular action (Garner, 2004). In a legal context, it is often used to emphasize that the procedure through which a judgment is made must be consistent, fair, independent, and transparent (Knoepfel, 2019, p.4). While MOCI Regulation No. 5/2020 outlines the expected content moderation action by private ESOs, particularly in response to MOCI takedown requests, the regulation contains limited due process and accountability provisions.

In the context of takedown requests by MOCI, due process should mean granting ESOs and internet users a fair mechanism to protest, challenge, and/or question the requests. Unfortunately, the regulation is not equipped with provisions for appeal mechanisms or even transparency reports over the removal of content. Instead, it pushes for private ESOs to be responsive towards both the general public and government's requests under strict time limits without room to express disagreement. Without due process, content moderation risks becoming an unchecked, top-down mechanism that may violate the users' freedom of expression.

Responding to that situation, a co-regulation approach could play an important role. Co-regulation gives room for risk-sharing responsibility. That refers to the distribution of roles and responsibility between the government and the ESOs in governing the digital space to minimize the risk created by harmful content as well as violation of users' freedom of speech.

Responding to that situation, a co-regulation approach could play an important role. Co-regulation gives room for risk-sharing responsibility.

Co-regulation can also become the avenue for knowledge sharing pertaining to the types of prohibited content spread across online platforms. Finally, co-regulation will create a more adaptable online regulatory environment because it allows for continuous evaluation and feedback from both the MOCI's and the ESOs' ends.

Challenges in Defining “Prohibited Content” Underscores the Need for Due Process

The lack of due process becomes even more problematic in light of the vague definition of prohibited content in MOCI Regulation No. 5/2020. Prohibited content, per this regulation, is (1) in violation of laws and regulations, (2) creates a public disturbance and disrupts public order, and (3) posts or provides access to prohibited content (Article 9(4)b). Uncertainties arise in areas such as who gets to decide if content disrupts public order, the difference between urgent and non-urgent content, or how the public and ESOs interpret this phrase.

Article 9 (5) stresses that government ministries or agencies determine whether content is creating a public disturbance and disrupting public order based on existing laws and regulations. Only a limited role is left for ESOs in interpreting these loosely defined terms. Some fear that concentration of power in the executive branch could put at risk users' freedom of speech (SAFEnet, 2021b). When the power to prohibit content is heavily invested in government bodies, a lack of due process provisions for the public and ESOs to fight for their rights undermines the implementation of fair and democratic content moderation.

Defining prohibited content has never been an easy task, nor is it unique to Indonesia. The sheer volume and variety of internet content, combined with the many perceptions of what content causes harm, makes it impossible to neatly categorize content according to a rigid and broadly accepted definition. Vague classification of prohibited content may be misused to execute and justify political oppression, persecution, and to limit freedom of expression.

But problems that arise from too broad a definition do not imply that a narrowed definition is a panacea. Overly granular definition of prohibited content leaves little room to consider context and may merely arrive through a different avenue at equally repressive provisions. Myanmar's attempts to provide a specific definition has resulted in excessive bans on subjective content seen as “destroying unity” or “inappropriate to Myanmar's culture” (Panday, 2021).

Overly granular definition of prohibited content leaves little room to consider context and may merely arrive through a different avenue at equally repressive provisions.

Because content moderation is inherently subjective (Singh, 2019), even if lawmakers attempt to make the definition more detailed and objective, the moderator's

subjective judgement will always be used to apply the definition to potentially prohibited content. India has tried to flesh out its definition of forbidden content through the IT Guidelines for Intermediaries and Digital Media Ethics Code Rules 2021 (IT Rules 2021), but still ends up with phrases such as “creating material risk of harm to sovereignty, integrity, state security, or public order” that need subjective interpretation by moderators to be applied (Rodriguez et al., 2021).

In Indonesia, similar concerns have been expressed with regards to Government Regulation (GR) No. 71/2019 on Implementation of Electronic Transactions and Systems and MOCI Regulation No.5/2020, except they apply to a much wider base. If Germany’s NetzDG and India’s IT Rules 2021 limit application to platforms that meet certain thresholds, Indonesian content moderation applies to all sizes of online platforms via the all-encompassing ESOs definition. This means that content moderation using a vague definition and applied with lack of due process affects the widest digital range.

Since objective definitions with clear-cut application are impossible, MOCI must guarantee that content moderation enforcement relies on due process and rules that create accountability among regulators and regulated alike, protecting freedom of expression.

Increasing Global Demand for Due Process in Content Moderation

Indonesia diverges from the global norm of including an appeal mechanism in content moderation undertakings by failing to include one in MOCI Regulation No. 5/2020.

There are several policy trajectories worth to be considered by MOCI in this context. The UK government is drafting an Online Safety Bill which includes provision on appeal mechanisms for unfair content removal (Beck-Watt, 2021; Wessing, 2021). In addition, Canadian academics, for example, have included a Moderation Standards Council meant to oversee moderation actions, facilitate appeal requests, and improve transparency to the public as one of their recommendations to the Canadian government (De Streel et al., 2020). Rather than appointing itself as an implementing and supervising institution in content moderation, MOCI could consider introducing such a council to improve the independence and accountability of content moderation governance in the country.

In addition to an appeal mechanism, transparency reporting has become part of the accountability mechanism in Malaysia, UK, France, and Germany (UK Ministry of State for Digital and Culture, 2020; Malaysian Communications and Multimedia Commission, 2021; De Streel et al., 2020). Transparency reporting refers to the issuance of reports that explain content moderation actions, including the reasons behind the decisions. The reports are issued by the moderator—the government and the ESOs—and are made available to the public.

In MOCI Regulation No. 5/2020, transparency reporting is not required by the ESOs or MOCI. However, some ESOs, like Facebook and Google, voluntarily publish their own reports (Facebook Transparency, 2020; Google, 2021). MOCI has also followed this step by including “actions on negative contents” in its annual reports, though it does not include details such as the reasoning behind MOCI’s decision, who was involved in the process, how the mechanisms work, or analysis of the context.

Limited Content Moderation Action

The loose definition plus the lack of clear-cut accountability mechanisms in the regulation creates a risk of over-enforcement, especially because MOCI Regulation No. 5/2020 almost explicitly limits its interpretation of content moderation to content removal.

“Limiting the regulatory conception of content moderation to content removal is problematic. Takedown requests from MOCI leaves only two extreme choices—to delete or to keep the reported content.”

Limiting the regulatory conception of content moderation to content removal is problematic. Takedown requests from MOCI leaves only two extreme choices—to delete or to keep the reported content. While the deleting may lead to an infringement of freedom of speech, keeping the content exposes the ESOs to sanctions from MOCI. Delete-or-keep is a rather stunted approach. In practice, UGC platform giants like Facebook, Twitter, and YouTube use a range of options, including depublishing, demoting, downranking, demonetizing, and warning users about potentially sensitive and harmful posts. This is useful to mitigate the risk of the platform going into either extreme. Also, applying graduated sanctions based on the severity of the violations may help to see if the culprits are willing to correct their behavior without platforms risking a violation of freedom of speech.

Content moderation actions adopted by UGC platforms giants like Facebook, Twitter, and YouTube warn users about potentially sensitive content, require users to log in for age verification prior to accessing certain materials, or flag content with potentially disputed status. For example, Twitter decided to flag and mark disputed claims in the 2020 U.S election rather than performing a mass takedown of tweets in its platform. These proven tools should be considered options beyond takedown requests that can be incorporated into MOCI regulatory policy.

THE IMPACT OF THE LACK OF DUE PROCESS IN MOCI REGULATION NO. 5/2020 ON CONTENT MODERATION MECHANISM FOR ESOS

ESOs' Dilemma in Responding to MOCI's Takedown Requests Within a Restricted Time Span

The lack of due process, narrowly defined content moderation actions, and vague definition of prohibited content creates challenges for ESOs. Since ESOs are only given 24 hours to remove non-urgent materials and four hours to remove urgent content, a proper assessment of any report that results in a takedown request is likely infeasible. This is especially true when materials are reported based on the qualitative grounds of causing a public disturbance or disrupting public order, which are difficult to measure. This limits ESOs to two equally difficult options—to hastily take down the content, potentially being accused of violating individual rights, or to dismiss MOCI's request and face up to three administrative fines before their internet access is blocked (Article 9 (6), Article 15 (7), Article 15 (12)).

The four or 24 hour timeline for takedowns is amongst the most stringent timelines globally. Even Germany's NetzDG, which is often considered as the prototype for excessive state censorship, allows 24 hours to seven days for the ESOs to assess, verify, and decide on their actions after receiving a request (Setiawan, 2021).

Similarly, the EU's E-commerce Directive, a foundational legal framework for internet services in the EU, doesn't require legislators in each member state to establish a specific time period. The rationale behind this is to uphold the principle of proportionality and reasonable decision-making that allow for flexibility in justified situations (De Streel et al., 2020). A strict time frame to comply threatens to create a "disproportionate burden on online platforms" and may risk encouraging over-enforcement that infringes on freedom of expression (De Street et al., 2020, p. 91)—concerns that are also relevant in Indonesia's case.

Action against allegedly prohibited content must take into account context, its potential impact on society, the size of the platform, and the nature of the service (De Streel, et al., 2020, p. 9). Seemingly similar content may have different effects depending on their context, intention, and where they are posted. There is also a global trend, especially strong in the EU, toward encouraging user participation in the notice-and-takedown system as opposed to relying on government requests (De Streel, et al., 2020).

Action against allegedly prohibited content must take into account context, its potential impact on society, the size of the platform, and the nature of the service.

France's Constitutional Council declared unconstitutional several provisions in the Avia Law, a bill targeting online hate speech that made the French government the authority for content

moderation. When the court struck these provisions down, authority for content moderation was returned to the platforms. The Council also stresses that it is up to the platforms to assess contents in their networks to allow for a more context-based judgment (Mchangama, 2021).

User Generated Content Platform Liability Under Self-initiated Content Moderation

The liability of private ESOs for content posted on their platforms, especially in the case of UGC platforms, is also an issue. The mandate for the UGC platforms to perform self-initiated content moderation comes with burdensome punishments for non-compliance—fines, access blocking (Article 10(5)) and holding them liable for hosting prohibited content (Article 11)—but with few rewards.

Article 11 of MOCI Regulation No. 5/2020 provides little reassurance for UGC platforms even if they manage to remove prohibited content. The article states that UGC ESOs that have undertaken self-initiated content moderation, including taking down prohibited content, “may be exempted” (*dapat dibebaskan*) from legal liability for hosting prohibited content. The phrase “may be” suggests that there is no guarantee that ESOs protect themselves through compliance.

Strengthening the Safe Harbor for ESOs to Strengthen Public-Private Collaboration

Providing a full safe harbor for online platforms—that is, protection from legal liability when certain conditions are fulfilled—plays a central role in successful content moderation, especially when MOCI has limited control over content on ESOs’ platforms. While it is reasonable to expect online platforms to prevent the spread of harmful content that is under their control, there should be legal protection for platforms that have managed to do this. Japan’s Provider Liability Law, India’s Information Technology (IT) Act, and Cambodia’s Law on E-Commerce all require ESOs to, where possible, prevent the distribution of harmful content on their platforms and reward those that do with full legal safe harbor (Provider Liability Law, 2001; IT Act, 2000; Law on E-Commerce, 2020).

MOCI Regulation No.5/2020 should therefore provide full safe harbor to ESOs by amending the phrase “may be exempted” (*dapat dibebaskan*) in Article 11 to “will be exempted” (*akan dibebaskan*) when referring to UGC ESOs that successfully self-initiated their content moderation.

RECOMMENDATIONS

Provisions in MOCI Regulation No. 5/2020 could violate users' freedom of speech due to the absence of due process, which is exacerbated by the vague definition of prohibited content and the narrow interpretation of what constitutes content moderation. Several policy changes would improve the regulation to ensure fair content moderation that protects users from harm as well as their freedom of expression:

MOCI should adopt a co-regulation approach

MOCI should engage with the private ESOs, civil society organizations, and other relevant stakeholders in the decision making process and attempt to adopt a regulatory approach that stresses shared responsibility for policy creation, implementation, and enforcement—also known as co-regulation (Aprilianti & Dina, 2021). This approach would help MOCI amend the provisions in MOCI Regulation No. 5/2020 and develop a content moderation guideline based on responsibility- and risk-sharing perspective.

Using the co-regulation approach, MOCI needs to amend several articles and accommodate the following:

MOCI Regulation No. 5/2020 should widen its interpretation on content moderation to give moderators more tools than merely content removal

MOCI Regulation No. 5/2020 almost explicitly limits content moderation to the removal of content from platforms. Content moderation should be understood more broadly as actions that limit the spread of content, including depublishing, demoting, downranking, and demonetizing, and warnings for users about possibly sensitive and harmful posts.

When content moderation is limited to content deletion, MOCI and ESOs are both more likely to violate freedom of expression. MOCI should allow private ESOs to use more tools to prevent the spread of harmful content while better balancing concerns about free expression.

MOCI Regulation No. 5/2020 should grant immunity to ESOs that have met their obligations under the regulation

MOCI should grant legal immunity for UGC ESOs that have performed content moderation in line with MOCI Regulation No. 5/2020. The term "may be exempted from legal liabilities" in the Article 11 needs to be amended into "will be exempted from legal liabilities" when referring to the UGC ESOs that have performed their obligations.

MOCI Regulation No. 5/2020 should include due process for ESOs and users, a reasonable time span for ESOs to comply, and an appeal mechanism

Rather than an environment of legal uncertainty and arbitrary decision making, MOCI Regulation No. 5/2020 should establish a due process for content moderation that allows ESOs and users to express their disagreement with MOCI's content removal requests and ensures that content moderation is transparent, independent, and fair.

MOCI should develop an appeal mechanism that allows ESOs and users to contest MOCI's takedown requests. This is to ensure that takedown requests are appropriately assessed within the corridor of freedom of expression principles. As a part of this, establishing an independent oversight board that includes representatives from MOCI, private ESOs and civil society to resolve interpretation disputes over the contents is also worth considering.

The timeframe in which ESOs are expected to comply with a MOCI takedown order should be made flexible with proportionate measures. Action against allegedly prohibited content must take into account context, the potential impact of the content on society, the size of the platform, and the nature of the service. More time is required for ESOs to comprehensively assess these factors against their community standards and relevant laws and regulations before they can appropriately act on them.

REFERENCES

- Adjie, M. (2020). *Internet ban during Papua antiracist unrest ruled unlawful*. Retrieved from The Jakarta Post: <https://www.thejakartapost.com/news/2020/06/03/internet-ban-during-antiracism-unrest-in-papua-deemed-illegal.html>
- Aprilianti, I., & Dina, S. (2021). *Co-regulating the Indonesian Digital Economy*. Retrieved from Center for Indonesian Policy Studies: https://c95e5d29-0df6-4d6f-8801-1d6926c32107.usrfiles.com/ugd/c95e5d_80051d5887a24aadb9a7c9f8f1cafc05.pdf
- Article19. (2017). *Joint Declaration on Freedom of Expression and “Fake News”; Disinformation and Propaganda*. Retrieved from Article19: <https://www.article19.org/resources/joint-declaration-on-freedom-of-expression-and-fake-news-disinformation-and-propaganda/>
- Article19. (2018). *Twitter Rules and Policies, Legal Analysis, August 2018*. Retrieved from Article19: <https://www.article19.org/wp-content/uploads/2018/09/Twitter-Rules-and-Policies-August-2018-.pdf>
- Barret, P. (2020). *Who Moderates the Social Media Giants? NYU Stern Center for Business and Human Rights*. Retrieved from: https://bhr.stern.nyu.edu/tech-content-moderation-june-2020?_ga=2.162404156.1232000728.1621356889-1606801777.1620065440. NYU Stern Center for Business and Human Rights.
- Beck-Watt, S. (2021). Federal Government Provides New Details of the Upcoming “Online Harms” Legislation and Regulator. Retrieved from Marks&Clerk: <https://www.marks-clerk.com/insights/federal-government-provides-new-details-of-the-upcoming-online-harms-legislation-and-regulator/>
- Caster, M. (2021). *Big tech needs a reset on Chinese censorship*. Retrieved from Reuters. <https://news.trust.org/item/20210607110849-zrpex>
- De Streel, A., D. E., Jacquemin, H., Ledger, M., Michel, A., Innessi, A., . . . Ustowski, D. (2020). *Online Platforms’ Moderation of Illegal Content Online*. Retrieved from European Parliament: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU\(2020\)652718_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU(2020)652718_EN.pdf)
- Deutscher Bundestag. (2020). *Gesetz gegen Rechtsextremismus und Hasskriminalität beschlossen*. Retrieved from Deutscher Bundestag: <https://www.bundestag.de/dokumente/textarchiv/2020/kw25-de-rechtsextremismus-701104>
- Facebook Transparency. (2020). *Content Restrictions Based on Local Law*. Retrieved from Facebook Transparency: <https://transparency.fb.com/data/content-restrictions>
- Feeney, M., & Duffield, W. (2020). *A Year of Content Moderation and Section 230*. Retrieved from CATO Institute: <https://www.cato.org/blog/year-content-moderation-section-230>
- Garner, B. A. (2004). *Black’s law dictionary, 8th Edition*.
- Gillespie, T. (2018). *Custodians of the Internet*. Yale University Press.
- Google. (2021). *Government Requests to Remove Content*. Retrieved from Google Transparency Report: <https://transparencyreport.google.com/government-removals/overview>
- IT Act. (2000). *The Information Technology Act, 2000*. Retrieved from Parliament of India: <https://www.indiacode.nic.in/bitstream/123456789/1999/3/A2000-21.pdf>
- Kerry, C. (2021). *Section 230 Reform Deserves Careful and Focus Consideration*. Brookings Institute. Retrieved from Brookings Institute: <https://www.brookings.edu/blog/techtank/2021/05/14/section-230-reform-deserves-careful-and-focused-consideration/>

Knoepfel, E. (2019). *The need for due process in Notice-and-takedown based content moderation on social media platforms*. Dickson Poon School of Law, King's College London.

Law on E-commerce. (2020). *Law on E-commerce*. Retrieved from Cambodia Ministry of Commerce: https://data.opendevdevelopmentcambodia.net/en/laws_record/law-on-e-commerce

Lomba, N., Navarra, C., & Fernandes, M. (2021). *Combating gender-based violence: Cyber violence*. Retrieved from European Parliament: [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662621/EPRS_STU\(2021\)662621_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662621/EPRS_STU(2021)662621_EN.pdf)

Malaysian Communications and Multimedia Commission. (2021). *MCMC Annual Report 2018*. Retrieved from Malaysian Communications and Multimedia Commission: https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/MCMC-2018_ENG.

Mchangama, J. (2021). *Rushing to Judgment: Examining Government Mandated Content Moderation*. Retrieved from Law Fare: <https://www.lawfareblog.com/rushing-judgment-examining-government-mandated-content-moderation>

Morar, D., & dos Santos, B. (2020). *Online Content Moderation, Lessons from Outside the US*. Retrieved from Brookings Institute: <https://www.brookings.edu/blog/techtank/2020/06/17/online-content-moderation-lessons-from-outside-the-u-s/>

Panday, J. (2021). *Regulation of Digital Platforms in Asia*. Retrieved from Internet Governance Project: <https://www.internetgovernance.org/2021/03/18/regulation-of-digital-platforms-in-asia/>

Provider Liability Law. (2001). *Act on the Limitation of Liability for Damages of Specified Telecommunications Service Providers and the Right to Demand Disclosure of Identification Information of the Senders Act No. 137 of 2001*. Retrieved from http://www.unesco.org/culture/pdf/anti-piracy/Japan/Jp_%20LimitLiability_Telecom_en

Rodriguez, K., Mathew, S., & Schmon, C. (2021). *India's Strict Rules For Online Intermediaries Undermine Freedom of Expression*. Retrieved from Electronic Frontier Foundation: <https://www.eff.org/deeplinks/2021/04/indias-strict-rules-online-intermediaries-undermine-freedom-expression>

SAFEnet. (2021a). *Permenkominfo 5/2020: Peraturan Moderasi Daring yang Diterapkan di Indonesia Mungkin Yang Paling Represif di Dunia*. Retrieved from SAFEnet: <https://id.safenet.or.id/2021/02/pm-kominfo-5-2020-peraturan-moderasi-daring-yang-diterapkan-di-indonesia-mungkin-yang-paling-represif-di-dunia/>

SAFEnet. (2021b). *Kertas Posisi Analisis Atas Permenkominfo No.5 Tahun 2020 Tentang Penyelenggaraan Sistem Elektronik Lingkup Privat*. Retrieved from SAFEnet: <https://id.safenet.or.id/2021/05/kertas-posisi-analisis-atas-permenkominfo-no-5-tahun-2020-tentang-penyelenggaraan-sistem-elektronik-lingkup-privat/>

Setiawan, I. (2021). *Who is Responsible for User-Generated Content on Digital Platforms in Indonesia?* Retrieved from Center for Indonesian Policy Studies: <https://www.cips-indonesia.org/post/policy-brief-who-is-responsible-for-user-generated-content-on-digital-platforms-in-indonesia>

Singh, S. (2019). *Everything in Moderation*. Open Technology Institute. Retrieved from <https://www.newamerica.org/oti/reports/everything-moderation-analysis-how-internet-platforms-are-using-artificial-intelligence-moderate-user-generated-content/>

Statista. (2020). *Media usage in an internet minute as of August 2020*. Retrieved from Statista: <https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/>

Statistics Indonesia. (2019). *Statistik Telekomunikasi Indonesia*. Statistics Indonesia.

Statistics Indonesia. (2020). *Survei Sosial Ekonomi Nasional (SUSENAS) 2015-2019*. Statistics Indonesia.

Statistics Indonesia. (2021). *Hasil Sensus Penduduk 2020*. Retrieved from Statistics Indonesia: [https://www.bps.go.id/pressrelease/2021/01/21/1854/hasil-sensus-penduduk-2020.html#:~:text=Hasil%20Sensus%20Penduduk%20\(SP2020\)%20pada,sebesar%20270%2C20%20juta%20jiwa](https://www.bps.go.id/pressrelease/2021/01/21/1854/hasil-sensus-penduduk-2020.html#:~:text=Hasil%20Sensus%20Penduduk%20(SP2020)%20pada,sebesar%20270%2C20%20juta%20jiwa)

UK Ministry of State for Digital and Culture. (2020). *UK White Paper: The Government Report on Transparency Reporting in relation to Online Harms*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/944320/The_Government_Report_on_Transparency_Reporting_in_relation_to_Online_Harms.pdf

Vogels, E. (2021). *The State of Online Harassment*. Retrieved from Pew Research Center: <https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment>

We Are Social & Hootsuite. (2021). *Digital 2021: Indonesia Report*. Retrieved from Data Reportal: <https://datareportal.com/reports/digital-2021-indonesia>

Wessing, T. (2021). *UK government sets out final approach to regulating online harms*. <https://www.lexology.com/library/detail.aspx?g=bc5e9b84-c70f-4eaf-bcd5-65e510603b25>

ABOUT AUTHORS

Pingkan Audrine is a Researcher at Center for Indonesian Policy Studies, focusing on the topic of Community Livelihood. She obtained her bachelor degree in Political Science from Parahyangan Catholic University. Prior to joining CIPS, Pingkan had experiences working in national broadcasting radio, international office at higher education institution, and Office of the UN Resident Coordinator in Indonesia.

Indra Setiawan specializes on the topics of agriculture and the digital economy in CIPS. Previously, he worked in a consulting firm as a research analyst on infrastructure development and the energy and natural resources policy. He graduated with a Bachelor's degree in International Relations at Universitas Diponegoro in 2018 and he is an alumnus of the CIPS Emerging Policy Leaders Program (EPLP) 2020.

Thomas Dewaranu holds a master's degree in public policy from the Australian National University and a bachelor's degree in law from Universitas Indonesia. His research interests cover rural development and poverty reduction. Prior to joining CIPS, he worked in a commercial law firm in Jakarta, providing legal services to local and multinational companies.

JOIN OUR SUPPORTERS CIRCLES

Through our Supporters Circles, you, alongside hundreds of others, enable us to conduct our policy research and advocacy work to bring greater prosperity to millions in Indonesia.

Those in our Supporters Circles get the opportunity to engage in the work of CIPS on a deeper level. Supporters enjoy:

- Invitation to CIPS' annual Gala Dinner
- Exclusive Supporters-only briefings by CIPS leadership
- Priority booking at CIPS-hosted events
- Personal (Monthly/Quarterly) Supporters-only update emails and videos
- Free hard copy of any CIPS publication upon request



For more info, please contact anthea.haryoko@cips-indonesia.org.



Scan to join

ABOUT THE CENTER FOR INDONESIAN POLICY STUDIES

Center for Indonesian Policy Studies (CIPS) is a strictly non-partisan and non-profit think tank providing policy analysis and practical policy recommendations to decision-makers within Indonesia's legislative and executive branches of government.

CIPS promotes social and economic reforms that are based on the belief that only civil, political, and economic freedom allows Indonesia to prosper. We are financially supported by donors and philanthropists who appreciate the independence of our analysis.

KEY FOCUS AREAS:

Food Security & Agriculture: To enable low-income Indonesian consumers to access more affordable and quality staple food items, CIPS advocates for policies that break down the barriers for the private sector to openly operate in the food and agriculture sector.


Education Policy: The future of Indonesia's human capital need to be prepared with skills and knowledge relevant to the 21st century. CIPS advocates for policies that drive a climate of healthy competition amongst education providers. Such competition will drive providers to constantly strive to innovate and improve education quality for the children and parents they serve. In particular, CIPS focuses on the improvement of operational and financial sustainability of low-cost private schools who serve the poor.


Community Livelihood: CIPS believes that strong communities provide a nurturing environment for individuals and their families. They must have the rights and capacities to own and manage their local resources and to ensure healthy and sound living conditions for the development and prosperity of the community.


www.cips-indonesia.org

 facebook.com/cips.indonesia

 [@cips_id](https://twitter.com/cips_id)

 [@cips_id](https://www.instagram.com/cips_id)

 [Center for Indonesian Policy Studies](https://www.linkedin.com/company/center-for-indonesian-policy-studies)

 [Center for Indonesian Policy Studies](https://www.youtube.com/channel/UC...)

Jalan Terogong Raya No. 6B
Cilandak, Jakarta Selatan 12430
Indonesia