

# KISIP 2024

Konferensi Ilmu Sosial dan Ilmu Politik



CSIS  
INDONESIA

Google

Research Paper

## Content Moderation in Southeast Asia General Elections: Case of Hate Speech in Indonesia, Malaysia and the Philippines

---

*Panel 4*

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

# M. Hafizh Nabiyyin

**Southeast Asia Freedom of Expression Network (SAFEnet),  
Indonesia**

✉ [hafizh@safenet.or.id](mailto:hafizh@safenet.or.id)

Hafizh Nabiyyin is the Head of Freedom of Expression Division at SAFEnet. He holds a Bachelor of International Relations degree from Universitas Potensi Utama, Medan.

# Balqis Zakiyyah Qonita

**Southeast Asia Freedom of Expression Network (SAFEnet),  
Indonesia**

✉ [balqis@safenet.or.id](mailto:balqis@safenet.or.id)

---

*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 It has not been peer-reviewed or been subject to internal review by SAIL, CSIS, or Google.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

Diverse forms of state and government systems are characteristic of the Southeast Asian region. Based on various studies, three countries that are considered as champions of electoral democracy in this region are Indonesia, the Philippines and Malaysia. The presence of social media as a public space that allows for political debate creates new challenges for the democratic resilience of these three countries. At least, the 2019 Presidential Election in Indonesia, the 2022 General Election in Malaysia, and the 2022 Presidential Election in the Philippines show a similar trend: the spread of hate speech that is harmful to minority groups in a democratic country. Continuing the 5-year cycle, Indonesia will again hold presidential elections in 2024. Therefore, this study aims to compare regulations related to content moderation implemented by the governments of Indonesia, Malaysia and the Philippines ahead of the election to prevent the spread of content containing hate speech to minority groups and look at its effectiveness. This research is empirical research that uses qualitative methods and a comparative approach. Data collection techniques are through literature study and interviews. The data were obtained from primary sources (informants, laws and regulations) and secondary sources (books, electronic journals, research reports, and electronic pages). This research concludes that the approach taken by these countries is still punitive and tends to ignore respect for human rights. The results are not effective, hate speech continues to haunt minority groups in cyberspace. This research produces recommendations at two levels: national and regional. At the national level, this research recommends a form of multi-stakeholder content regulation, together with civil society, academics, social media companies and other relevant actors for each country. At the regional level, this research recommends trilateral cooperation "Maphilindo" to create content moderation policies and practices that are in line with international human rights standards and local contexts. Cooperation can be done in carrying out prevention, such as increasing digital literacy and negotiating with social media platforms.

**Keywords:** *Content moderation, general election, Southeast Asia, hate speech, democracy*

## BACKGROUND

Southeast Asia is a sub-region in the Asia Pacific region which is inhabited by more than 689 million people<sup>1</sup>. This region is very diverse, both in terms of culture, language, ideology, religion, economic progress, state form, and government system. There are at least 11 countries that inhabit this region, namely Indonesia, Thailand, Malaysia, Singapore, the Philippines, Vietnam, Cambodia, Laos, Brunei Darussalam, Myanmar and Timor Leste. Apart from Timor Leste, ten other countries have been gathered into the regional organization Association of Southeast Asia Nations (ASEAN). ASEAN itself was founded by the first five countries in 1967 and has gradually succeeded in embracing other countries in the region.

Since it was first established, ASEAN has transformed from an organization whose focus was limited to cooperation in the economic, security and defense sectors to also touching on democracy and human rights issues. This can at least be seen from the progress in recognizing democracy and human rights in various ASEAN documents, especially those created after the economic crisis that hit the region in 1998. Some of these documents include the ASEAN Charter (2007) and the ASEAN Human Rights Declaration (2012).

Although normatively democracy and human rights have progressed in ASEAN institutions, developments are more dynamic and fluctuating in its member countries. Croissant (2022) divides 11 countries in the Southeast Asia region into three regime forms. Cambodia, Singapore, and – according to several studies – the Philippines are examples of “electoral authoritarianism”. The second group, “closed autocracies”, namely regimes that avoid competition in elections, consists of Brunei Darussalam, Laos, Vietnam, Myanmar and Thailand. Laos and Vietnam openly claim to be communist party states, Brunei is an absolute monarchy, while Myanmar and Thailand are directly or indirectly ruled by the military. The third category is “electoral democracy,” including Indonesia, Malaysia, Timor Leste, and, perhaps, the Philippines.<sup>2</sup>

Based on the Croissant typology above, Indonesia, Malaysia and the Philippines are the three representatives of electoral democracy in ASEAN (Timor Leste is not yet an official member). The V-Dem report in 2022 shows Indonesia as the best country in the electoral democracy index, followed by Malaysia and the Philippines<sup>3</sup>. In the context of freedom, the Freedom in the World report shows the three countries as champions of freedom in the region. Indonesia is in first place, followed by the Philippines and Malaysia<sup>4</sup>. From the 10 ASEAN member countries, only three countries are said to be partly free, while the rest are said to be not free. Freedom on the Net report is not much different, the Philippines and Malaysia are the champions, followed by Indonesia.<sup>5</sup>

Taking into account the various reports above, Freedom on the Net is interesting to look into further. The reason is that recently the discourse on digital

---

<sup>1</sup> <https://www.worldometers.info/world-population/south-eastern-asia-population/>

<sup>2</sup> [https://link.springer.com/chapter/10.1007/978-3-031-05114-2\\_1](https://link.springer.com/chapter/10.1007/978-3-031-05114-2_1)

<sup>3</sup> [https://v-dem.net/data\\_analysis/CountryGraph/](https://v-dem.net/data_analysis/CountryGraph/)

<sup>4</sup> <https://freedomhouse.org/explore-the-map?type=fiw&year=2023>

<sup>5</sup> <https://freedomhouse.org/explore-the-map?type=fofn&year=2023>

transformation has continued to emerge in this region. By 2022, there at least 400 million internet users in Southeast Asia with a penetration rate of more than 70 percent in all countries except Laos, Myanmar and Timor Leste<sup>6</sup>. This progress cannot be separated from the leaders of countries in the region as at least stated in the ASEAN Leaders' Statement on Advancing Digital Transformation in ASEAN which was delivered at the 2021 ASEAN Summit. Apart from that, the agenda for increasing internet access is also stated in the Master Plan on ASEAN Connectivity 2025 and ASEAN Digital Master Plan 2025.

Even though it is quite progressive in introducing digital transformation, the impact of digitalization on freedom, which is one of the foundations of democracy, has not been widely considered by ASEAN policy makers. In fact, Freedom on the Net depicts the emergence of a trend towards digital authoritarianism in several countries in the region. This is counterproductive to the progress the region has achieved in promoting democracy and human rights. Two of the indicators used in the report are the state's role in limiting content and ensuring a safe online environment for vulnerable groups (LGBTIQ, ethnic minorities, religion, etc.) from online intimidation and harassment.

Hate speech, especially during elections, has become a global trend. The Southeast Asian region is not immune to this trend. The elections in three Southeast Asian democracy champions – Indonesia, Malaysia and the Philippines – in the past 5 years are a recent example. Paladino (2018) states that in Southeast Asia, there is manipulation of behavior at the grassroots level caused by the spread of hate speech and fake news by charismatic leaders and groups with the aim of attracting supporters for long-standing communal conflicts and threatening the consolidation of democracy.<sup>7</sup>

Bhat & Banaji (2022) highlight the impact of this hate speech on minority groups. According to him, inequality and prejudice against caste, class, race, disability, gender, sexuality and religion from real life can be widespread on social media, while social media can also be used to build and sustain racial, caste, gender, sexual and class identities.<sup>8</sup> Based on studies they conducted in four countries, namely India, Myanmar, Brazil, and the United Kingdom (England), minority groups of religion, ethnicity, race, sexuality, class, caste, disability, body image, and those who are vocal about their rights become among the targets of hate speech perpetrators.

The spread of hate speech, especially towards minority groups, has an impact on the progress of democracy and human rights in the region. Considering democracy, its amplification during elections can cause societal polarization and damage the integrity of elections. Meanwhile, from a human rights perspective, it can take away the right to feel safe for minority groups in the digital space. Therefore, state intervention is needed to ensure that social media spaces are free from hate speech, especially during elections. In this case, the state could be involved in regulating content moderation by social media companies.

This research aims to compare content moderation regulatory frameworks in Southeast Asia and produce concrete proposals that can be implemented in

---

<sup>6</sup> <https://www.statista.com/statistics/487965/internet-penetration-in-southeast-asian-countries/>

<sup>7</sup> [https://www.brookings.edu/wp-content/uploads/2018/07/FP\\_20180725\\_se\\_asia\\_social\\_media.pdf](https://www.brookings.edu/wp-content/uploads/2018/07/FP_20180725_se_asia_social_media.pdf)

<sup>8</sup> Social Media and Hate

order to save democracy in the region. Researchers limit the scope of research to three democracy champion countries in ASEAN, namely: Indonesia, Malaysia and the Philippines. The limited study objects are caused by limited time and resources that researchers have. The Philippines was chosen over Timor Leste because researchers saw its greater influence on other countries in the region. In addition, this research only highlights hate speech, not dis/misinformation, bullying, or other harmful content – which also often overlap. Without minimizing the impact of this dangerous content, the author believes that hate speech needs special attention because of its impact which can extend to communal conflict.

## FINDINGS

### Hate Speech on Southeast Asia General Elections

Hate speech is a form of verbal, written or behavioral communication that attacks or uses pejorative or discriminatory language against a person or group because of their religious background, ethnicity, nationality, race, skin color, descent, gender or other identity factors.<sup>9</sup> Hate speech (along with disinformation and propaganda) is considered to be one of the factors causing the rise of far-right groups in elections in various parts of the world. The victories of Donald Trump in the United States in 2016, Viktor Orban in Hungary in 2018 and 2022, and Jair Bolsonaro in Brazil in 2019, are popular examples of the correlation between the use of hate speech as a political campaign tool and the victory of right-wing politicians.

In Southeast Asia, the Asia Center (2020) classifies four forms of hate speech: (1) Hate speech related to ethnicity and religion; (2) Hate speech against foreign nationals, migrant workers and refugees; (3) Hate speech based on political values and ideology; and (4) Hate speech against sexual minorities. This research also states that the number of incidents of hate speech against sexual minorities increases during the election season, where sexual minority groups are targeted by political campaigns.<sup>10</sup>

Indonesia is the largest country in the region with a population of 277 million people.<sup>11</sup> With the largest population, internet penetration in Indonesia still reaches 76.5%.<sup>12</sup> Of this figure, 167 million people or 60.4% of the entire population are social media users. YouTube is the most popular social media in Indonesia with 139 million users<sup>13</sup>.

In 2019, for the fourth time, Indonesia held democratic general elections. At this event, elections were held simultaneously. Voters must cast five ballot papers at once: president and vice president, members of the People's Representative

---

<sup>9</sup>[https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech\\_Guidance%20on%20Addressing%20in%20field.pdf](https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech_Guidance%20on%20Addressing%20in%20field.pdf)

<sup>10</sup> [https://drive.google.com/file/d/1qdk70SedPXjuOkYWFUAVSsTKYKaIP78/view?usp=share\\_link](https://drive.google.com/file/d/1qdk70SedPXjuOkYWFUAVSsTKYKaIP78/view?usp=share_link)

<sup>11</sup> EDB Asia Partners Internet Report

<sup>12</sup> <https://www.statista.com/statistics/487965/internet-penetration-in-southeast-asian-countries/>

<sup>13</sup> <https://datareportal.com/reports/digital-2023-indonesia>

Council, members of the Regional Representative Council, members of the Provincial Regional People's Representative Council, and members of the Regency/City Regional People's Representative Council. The presidential election between incumbent Joko Widodo (Jokowi) and former general Prabowo Subianto received special attention. The fight between the two is a rematch of the 2014 presidential election, where Jokowi narrowly won over Prabowo.

That year's election was said to be the most brutal election in the history of democracy in Indonesia after Soeharto's authoritarian regime. Social media has become one of the main battlegrounds, where the official success teams and sympathizers of the two candidates attack each other using disinformation, propaganda and hate speech. During the fourth quarter of 2018 to the second quarter of 2019, Facebook recorded 11.9 million pieces of hate speech content. This time coincides with the campaign, election and announcement of the winner of the 2019 Indonesian Election.<sup>14</sup>

From the hate speech used, Chinese identity is often exploited to attack Jokowi with the narrative that he is of Chinese descent. The word "*aseng*" is one of the words often used to attack Jokowi. Furthermore, CSIS Indonesia conducted research by analyzing the word "*china*" on Twitter in 2019-2020. Looking at the historical context, this word is often used to exclude Indonesian Chinese. As a result, they found an average of 696 tweets per month in 2019 and 855 tweets per month in 2020.<sup>15</sup> The use of the word "Chinese" is juxtaposed with "communist", "indigenous", "country", "people", "corona", and "virus". Apart from targeting ethnicity, terms close to gender minorities such as "*bencong*", "*melambay*", and "sissy" are also used to attack political opponents.<sup>16</sup> Prabowo used anti-LGBT rhetoric as a populist weapon to mislead the public and attack Jokowi, claiming that Jokowi supported same-sex marriage.<sup>17</sup> In 2018 (campaign season), the local government of West Sumatra promised to criminalize LGBT after anti-LGBT demonstrations.

Malaysia has 34 million residents.<sup>18</sup> Even though it is small in population, internet penetration in Malaysia has reached 93.8% – the second highest in ASEAN after Brunei Darussalam.<sup>19</sup> Of this figure, 26.8 million people or 78.5% of the total population are social media users. Just like Indonesia, YouTube is the most popular social media in Malaysia with 23.9 million users.<sup>20</sup>

In 2022, the Malaysian Communications and Multimedia Commission reported that they received more than 1,700 complaints of hate speech cases related to race, religion and royal status (race, religion, royalties or 3R).<sup>21</sup> Zamri et

<sup>14</sup> <https://databoks.katadata.co.id/datapublish/2020/05/15/facebook-saring-jutaan-konten-ujaran-kebencian>

<sup>15</sup> [https://s3-csis-web.s3.ap-southeast-1.amazonaws.com/doc/Hate\\_Speech\\_Against\\_Chinese\\_Indonesians\\_2019\\_2020\\_\\_Policy\\_Brief.pdf?download=1](https://s3-csis-web.s3.ap-southeast-1.amazonaws.com/doc/Hate_Speech_Against_Chinese_Indonesians_2019_2020__Policy_Brief.pdf?download=1)

<sup>16</sup> <https://ejournal.ukm.my/mjc/article/view/39517>

<sup>17</sup> <https://drive.google.com/file/d/1qdk70SedPXjuOkYWFUAVSsTKYKalP78/view>

<sup>18</sup> EDB Asia Partners Internet Report

<sup>19</sup> <https://www.statista.com/statistics/487965/internet-penetration-in-southeast-asian-countries/>

<sup>20</sup> <https://datareportal.com/reports/digital-2023-malaysia>

<sup>21</sup> <https://drive.google.com/file/d/1a-iXd5LqnWHxDTWdDUEyikBAICiP82pz/view>

al (2020) found that racist sentiments were widely used, such as “china”, “keling”, and “meleis”. This terminology is an example of a racist term that targets the Chinese, Malay and Indian races. These findings are in line with social media monitoring carried out by the Center for Independent Journalists (CIJ) and analysis of TikTok carried out by Jalli during the 15th Malaysian Election.

CIJ is monitoring social media Facebook, Twitter, YouTube and Twitter during 20 October - 26 November 2022 (elections held on 19 November). Their monitoring focuses on hate speech against 3R (religion, royalty, race), gender and LGBTIQ, as well as refugees and migrants. The findings were that there were 99,563 uploads of hate speech, of which 39 were incitement to violence. Race-based narratives were the most numerous with 66,933 posts, followed by religion with 24,484 posts, royal status with 14,320 posts, gender and LGBTIQ with 7,596 posts, and comments targeting migrants and refugees with 3,819 posts.<sup>22</sup>

Race-based narratives were the focus of the study conducted by Jalli. After searching for uploads containing certain keywords in the two weeks before and after the 15th Election, he found that there were at least 373 videos with over 1000 viewers that contained hatred and propaganda. Jalli concluded that hate speech in Malay mostly targets non-Malay races, especially the Chinese community and conversely, Chinese language content focuses on the Malay and Muslim communities.<sup>23</sup> For example, there is a post with text that reads “*Haram Umat Melayu Islam undi PH (Pakatan Harapan)*”, considering that there is a Chinese-majority party, DAP, in the PH coalition. On the other hand, there were also TikTok users who uploaded videos containing DAP supporters' criticism of PAS (which is predominantly Muslim) by calling them “stupid Muslim ulama.” Apart from that, during the election the hashtags #13mei and #13mei1969 also appeared, referring to the conflict between the Malay and Chinese communities in Malaysia which was triggered by demonstrations rejecting the election results.

The population of the Philippines reaches 114 million, the second largest in ASEAN.<sup>24</sup> The internet penetration rate in this country has reached 91%.<sup>25</sup> Meanwhile, 72.5% of the population or the equivalent of 84.45 million people are active social media users. In this country, Facebook is the most popular social media platform with around 80.3 million users.<sup>26</sup>

In May 2022, the Philippines held presidential and vice presidential elections. Apart from that, elections were also held for members of the Senate, DPR, as well as regional heads and regional legislative members. As a result, Ferdinand “Bongbong” Marcos Jr. (son of former Philippine dictator Ferdinand Marcos) was elected president and Sara Duterte (daughter of incumbent president Rodrigo Duterte). Bongbong got 58.77% of the votes while Sara got 61.53%. These two populist figures were elected amidst hate speech that spread across Philippine social media.

---

<sup>22</sup> <https://drive.google.com/file/d/1u3yl29GVW2swLc1HNNkqWZFrUQszs6mU/view>

<sup>23</sup> <https://theconversation.com/how-tiktok-became-a-breeding-ground-for-hate-speech-in-the-latest-malaysia-general-election-200542>

<sup>24</sup> EDB Asia Partners Internet Report

<sup>25</sup> <https://www.statista.com/statistics/487965/internet-penetration-in-southeast-asian-countries/>

<sup>26</sup> <https://datareportal.com/reports/digital-2023-philippines>



Throughout January-May 2022, Facebook and Instagram noted that they had removed more than 6 million posts from Filipino users for violating their community guidelines<sup>27</sup>. As many as 670 thousand of them violated Meta policies because they contained hate speech, which is defined as "direct attacks on people based on protected identity." Researchers have not yet found out specifically which groups are the targets of hate speech. However, literature on the 2016 election found that hate speech circulating in the Philippines was based on race, physicality, gender, disability, religion, class and quality<sup>28</sup>.

Muslim groups, the majority of whom come from the Bangsamoro ethnicity, have been the targets of hate speech. During the siege of Marawi by ISIS-affiliated terrorist groups in May 2017, Philippine social media was filled with hate speech against Muslims. Words used include "traitors", "violent savages", "juramentado", "pirates", "assassins", "enslavers", "cruel", and "uncivilized"<sup>29</sup>.

Apart from that, women's groups were also targeted. Maria Ressa, a female journalist and CEO of online media Rappler, received numerous death and rape threats from trolls and Duterte supporters<sup>30</sup>. In the 2022 elections, Bongbong Marcos' opponent, Leni Robredo, also received a lot of hate speech. The attack pattern against Leni is red-tagging. Red-tagging can be interpreted as an action where state actors, especially law enforcement agencies, publicly and without basis label individuals, groups, or institutions as communist terrorist groups<sup>31</sup>. Research from Internews consistently finds a lot of sexual and degrading content attacking Robredo because of her appearance and gender<sup>32</sup>. For example, he was called a "weak leader" and sexually objectified in public forums. Apart from Leni, other public figures such as Ellen Tordesillas (Vera Files) have also been called "every inch a prostitute" and Leila de Lima (Senator) has been called "an immoral woman".

## Content Moderation Regulatory Framework in Southeast Asian Countries

The social media space must actually guarantee a balance in fulfilling human rights. The world is increasingly interconnected and as the fabric of society has essentially become more multicultural. A number of incidents in recent years, in various parts of the world, have brought renewed attention to the problem of hate speech. It should also be underlined that many conflicts around the world in recent decades also contain components of incitement to hatred against a nation, race or

---

<sup>27</sup> <https://thediomat.com/2022/08/facebook-instagram-purged-millions-of-posts-during-philippine-election/>

<sup>28</sup> <https://drive.google.com/file/d/1EVjRvJ9SiQR6nZGgz3WSacmO4vrpciwp/view>

<sup>29</sup> [https://gaamac.org/wp-content/uploads/2022/07/APSG-REPORT\\_FINAL.pdf#page=53](https://gaamac.org/wp-content/uploads/2022/07/APSG-REPORT_FINAL.pdf#page=53)

<sup>30</sup> <https://fma.ph/2022/06/29/understanding-hate-and-hate-speech-the-philippine-context/>

<sup>31</sup> <https://internews.org/wp-content/uploads/2023/01/Red-Tagging-in-the-Philippines.pdf>

<sup>32</sup> <https://internews.org/wp-content/uploads/2023/01/Red-Tagging-in-the-Philippines.pdf>

religion.<sup>33</sup> Seeing this problem, a content degradation system is needed that has objective limits to overcome digital pollution containing discrimination.

Content moderation is a way for platforms to reduce content that is harmful to other people. This mechanism is the organized practice of screening user-generated content (UGC) posted to Internet sites, social media, and other online outlets, to determine its suitability for a particular site, locality, or jurisdiction.<sup>34</sup> On social media, there are community standards for each platform that are established to remove dangerous and/or illegal content. Each platform carries out self-regulation to manage the timeline to suit the protection of users.

Each platform has various conditions. YouTube and Tiktok, as video and text-based platforms, prohibit posting content if it contains the aim of encouraging violence against individuals or groups. Meanwhile, Twitter regulates the prohibition of content that makes threats of violence against identifiable threats of violence. The threat of violence is declarative in nature to cause serious injury. Facebook will remove language that incites and facilitates serious violence. Meanwhile, Instagram will delete content that is believed to contain threats or hate speech.<sup>35</sup>

However, the weakness is that it cannot clearly distinguish between dangerous content and illegal content. There is also gray content that must be considered and strictly regulated. So, community guidelines alone are not enough. There is a need to understand local contexts. Resolution of content removal is still far from transparency and accountability. Platforms should be monitored in their actions so that the role of the state is needed to form regulations that balance the freedom of platforms to contract with users and encourage accountability for content moderation<sup>36</sup>. The state must guarantee protection within the border space so that all its citizens can obtain information well and safely.

The role of law in this condition is necessary. Everyone is obliged to act as society wishes, so that order is maintained as well as possible. The law that is formed must include various regulations that determine and regulate the relationships between people and each other, namely the rules of social life.<sup>37</sup>

## Indonesia

In the Indonesian constitutional framework, there is a guarantee of freedom of expression in Article 28E and the right to obtain information in Article 28F. However, this right is included in the derogable right category. Derogable rights are human rights that can be limited with the aim of maintaining public order and democracy. Article 28J of the 1945 Constitution states that everyone must comply with the restrictions contained in the law to guarantee recognition of respect for the rights and freedoms of other people.

The development of information and technology ultimately gave birth to *lex specialis* which regulates the transaction and transmission processes that occur in

---

<sup>33</sup> Ini lupa

<sup>34</sup> Sarah T Roberts, 'Content Moderation', 2017 <<https://escholarship.org/uc/item/7371c1hf>>.

<sup>35</sup> Sherly Haristya (Article 19), *Moderasi Konten Dan Pemangku Kepentingan Lokal Di Indonesia*, 2022.

<sup>36</sup> Moderasi konten dan pemangku kepentingan lokal di Indonesia, Article 19

<sup>37</sup> <http://ejournal.radenintan.ac.id/index.php/TAPi/article/view/1578/1305>

the virtual universe. The Electronic Information and Transaction Law (EIT Law) is a tool to regulate the broad and targeted use of information technology for the sake of creating an electronic society that always applies morals and ethics in all aspects of life.<sup>38</sup>

This ITE Law tries to accommodate restrictions on expression as mandated by the constitution. Article 28 paragraph 2 jo. Article 45 paragraph 2 states that everyone intentionally and without the right to disseminate information aimed at causing hatred or hostility of individuals and / or certain groups of people based on ethnicity, religion, race, and intergroup will be punished. This article gives a broad interpretation of protected objects because of the phrase between groups of many groups.

In the process of controlling action, Article 40 of the EIT Law legitimizes government action to regulate the electronic domain from all types of disturbances as a result of misuse of Electronic Information and Electronic Transactions which disrupt public order, which refers to the provisions of laws and regulations regarding. This provision gives the government the authority to terminate access and/or order PSE to do so. The derivative regulations also give the government great authority to terminate access because it requires PSE registration. Ministry of Communication and Information Regulation Number 5 of 2020 (MR5) allows the public to report prohibited content to the MoCI. The MoCI then verified whether the content was prohibited and submitted a request for deletion to PSE. After receiving a request from the MoCI, PSE has up to 24 hours to cut off access to content deemed not urgent, or four hours for those deemed urgent. Urgent content includes material related to terrorism, child pornography, and content that disturbs society or disturbs public order.<sup>39</sup>

At the end of 2023, the Government will also recodify the criminal code in the Criminal Code. Previously in the old Criminal Code, there was only an article on blasphemy. That any person in public who expresses feelings or commits acts that are essentially hostile, abusive, or blasphemous towards a particular religion can be punished..<sup>40</sup>

After the recodification, the government formulated the hate speech article in Number 1 of 2023 concerning the Criminal Code. Unlike the ITE Law, this new law mentions elements of hate speech more specifically. which expands the formulation with phrases between groups, in Article 243 paragraph (1) of the Criminal Code states that criminal provisions can be used on people who express feelings of hostility to result in violence against people or property to racial, national, ethnic, color, religious, gender, mental disability, or physical disability. The implementation of this article also becomes stricter because there is a description of groups that are protected from hatespeech and contain elements of material consequences.

As an election organizing institution, the Indonesian General Election Commission (KPU) makes rules during the campaign period during the election to

---

<sup>38</sup> file:///C:/Users/HP/Downloads/10279-28365-1-PB.pdf

<sup>39</sup>[https://c95e5d29-0df6-4d6f-8801-1d6926c32107.usrfiles.com/ugd/c95e5d\\_3d6ad81f6e1f4220919e453ddf9f3e8f.pdf](https://c95e5d29-0df6-4d6f-8801-1d6926c32107.usrfiles.com/ugd/c95e5d_3d6ad81f6e1f4220919e453ddf9f3e8f.pdf)

<sup>40</sup> KUHP pasal 156

create supportive and healthy competition. This regulation prohibits insulting a person, religion, ethnicity, race, group of candidates, and/or other election participants. This is stated in Article 69 paragraph (1) of KPU Regulation Number 32 of 2018. If it is proven that this act has occurred, sanctions will be imposed in accordance with the Law governing Elections and other laws and regulations as stated in Article 76. So, The moderation process if there is a violation of this article can refer to the ITE Law.

Election Supervisory Body (Bawaslu) Regulation Number 1 of 2018 Article 6 paragraph (1) letter c also includes provisions that Bawaslu, Provincial Bawaslu, and Regency/City Bawaslu carry out supervision over things that are prohibited in the implementation of the Campaign, one of which is insulting a person, religion, ethnicity, race, class, candidates, and/or other election participants.

## Malaysia

Meanwhile in the neighboring country, as the foundation of Malaysia's Federal Constitution, it includes guarantees for freedom of opinion and freedom of assembly in Article 10. However, in Article 8(2) there are also restrictions that do not allow discrimination based on religion, race, descent, place of birth, and gender.

Due to socio-cultural factors, racial politics has been identified as a concerning issue in Malaysia according to the Malaysia Racial Discrimination Report (2021). In 2021, a total of 53 racism and/or racial discrimination incidents were documented in Malaysia based on media reports and exploratory observation by Pusat KOMAS (2022).<sup>41</sup>

In terms of regulations, since 1948 there has been The Sedition Act which was formed to be a government tool for regulating and monitoring public discourse to prevent any public action that could be interpreted as 'sedition'. However, sedition in this law has a broad interpretation, which does not comply with the above standards governing restrictions on the right to freedom of expression. Seditious tendencies are defined in section 3(1) where one of the elements is an act that encourages hatred and enmity between races or classes.<sup>42</sup>

Malaysia and the government are also proposing the enactment of an Anti-Discrimination Bill, the National Harmony and Reconciliation Commission Bill, and the Racial and Religious Hatred Bill.<sup>43</sup> However, the bills have not progressed, and discussions have instead turned to amendments to existing legislation to address the issue of 'hate speech'.<sup>44</sup>

As the election organizer, the Malaysian Election Commission (EC) formulates regulations regarding the prohibition of hate speech during the General Election period. The Malaysian Election Offenses Act (1954) states that if any person who, before, during or after an election, directly or indirectly, by himself or

---

<sup>41</sup> Digital hate speech and othering: The construction of hate speech from Malaysian perspectives

<sup>42</sup>[https://www.ijbel.com/wp-content/uploads/2015/05/Law1\\_PAID\\_IJBEL\\_journal-IJBEL-vol.-6-Apr-2015\\_D1.pdf](https://www.ijbel.com/wp-content/uploads/2015/05/Law1_PAID_IJBEL_journal-IJBEL-vol.-6-Apr-2015_D1.pdf)

<sup>43</sup> CIJ

<sup>44</sup><https://www.article19.org/resources/malaysia-efforts-to-combat-hate-speech-should-not-trample-freedom-of-expression/>

by any other person on his behalf, does any act or makes any statement with a view or with a tendency to promote feelings of ill-will, discontent or hostility between persons of the same race or different races or of the same class or different classes of the population of Malaysia in order to induce any elector or voter to vote or refrain from voting at an election or to procure or endeavor to procure the election of any person will be imprisonment for a term not exceeding five years or to a fine not exceeding ten thousand ringgit or to both such imprisonment and fine.<sup>45</sup> The Act also prohibits direct or indirect undue influence through the use of force, force, or restraint, or by causing injury, damage, loss or loss.<sup>46</sup>

These two parts are specifically to ensure the freedom to exercise citizens' right to vote. However, these sections offer little security for non-citizens and marginalized groups which are often the target of hate speech as well as criminal investigations. To date, the Election Commission has yet to draft any specific guidelines on the use of social media during elections, leaving the responsibility of moderating in the hands of the platforms themselves or the Malaysian authorities.<sup>47</sup>

## The Philippines

Meanwhile, in the Philippines, the government has not passed any laws prohibiting hate speech, incitement to violence, and discrimination in general. There are no legal provisions against such kind of speech as jurisprudence on freedom of expression cases mainly focus on libel, defined in the public and malicious imputation of a discreditable act that tends to discredit or dishonor another person and which currently exists under the Revised Penal Code.

The Philippines does have a Cybercrime Prevention Act of 2012, for example, regulating internet crimes including online hate speech. However, this law does not specifically regulate hate speech and is often used to criminalize critical voices in the Philippines. The law even gives the Department of Justice the power to shut down websites and monitor online activity without a warrant. So, the Department of Justice has the right to moderate and it is feared that there will be no transparent legal process.<sup>48</sup>

Apart from that, there is a law against defamation and slander. During the 18th Congress, a bill was introduced in the House of Representatives aimed at defining, prohibiting, and punishing hate speech in the country. The bill was introduced by Muslim representatives in Congress following violent incidents in Christchurch and El Paso in which people were killed due to racist and anti-immigrant sentiment.<sup>49</sup>

---

<sup>45</sup> Election Offence Act 1954 Article 4A

<sup>46</sup> Election Offence Act section 9

<sup>47</sup> Laporan media Monitoring CIJ

<sup>48</sup> Philippines 'Cybercrime' <https://www.amnesty.org/en/latest/news/2012/10/philippines-cybercrime-law-threatens-freedom-expression-and-must-be-reviewed/>

<sup>49</sup> Understanding Hate and Hate Speech: The Philippine Context <https://fma.ph/2022/06/29/understanding-hate-and-hate-speech-the-philippine-context/>

## Hate Speech and the Right to Feeling Safe for Minority Groups

Cyber is indeed becoming an increasingly central forum for the exercise of a host of other rights as people increasingly look to the Internet to access information, form connections with others, and organize social life. In this context, it is crucial that States refrain from using cyber technology to violate human rights and must likewise refrain from interfering with or curtailing the enjoyment of human rights in cyberspace.<sup>50</sup> States are therefore obligated to “provide such safety in law and policy that will allow individuals to secure themselves online.” And the duty included to respect, to protect, and to fulfill the human right for everyone.

In a conventional context, every country has a basic right to security for its citizens. Indonesia in Article 28G paragraph (1) of the 1945 Constitution states that every citizen has the right to a sense of security and protection from the threat of fear to do or not do something. TAP MPR Number XVII / MPR998 Chapter VII classifies human rights and one of them is the right to security which also guarantees the same protection in Article 22. Article 25 also states that everyone should be free from torture or degrading treatment.

The Constitution of Philippines on section 11 states that The State values the dignity of every human person and guarantees full respect for human rights. Unfortunately, in the implementation stage, this sense of security still threatens certain groups in the online space and has implications for real life. In 2019, at that time riots were breaking out in Papua. Hoaxes circulated in Surabaya which ultimately resulted in hate speech leading to the siege of Papuan students in Surabaya. The impact of hatespeech ultimately shakes the peace in the physical space with threats of violence, expulsion, and the thickening of bad stigma.

During the election period in Indonesia in 2019, the Bawaslu Report showed that identity issues or hoaxes that emerged in that year's election, for example, Prabowo Christian, Jokowi Anti-Islam, Jokowi supporters of adultery and same-sex marriage, Prabowo supporters of polygamy, and so on (Kartini, 2019).

Joko Widodo, who is considered pro-China, causes hate narratives to Chinese groups to bring down Jokowi's image. Joko Widodo is labeled as creating policies that will facilitate investments from China and will send many workers from the country. In May 2019, after several people perished during the violent protest in Jakarta, messages circulated via social media suggesting the loss occurred because of police from China; the photo of masked police officers accompanied the claim with caption: China has sent security forces to Indonesia disguised as foreign workers (Chew,2019).<sup>51</sup>

Many LGBT groups also experience violence online with hate speech and this thickening of stigma is also threatened in real life. There is a widespread

---

<sup>50</sup> State Responsibility to Respect, Protect and Fulfill Human Rights Obligations in Cyberspace [https://jnslp.com/wp-content/uploads/2017/10/State-Responsibility-to-Respect\\_2.pdf](https://jnslp.com/wp-content/uploads/2017/10/State-Responsibility-to-Respect_2.pdf)

<sup>51</sup>Hatespeech in Southeast Asia <https://asiacentre.org/wp-content/uploads/2020/07/Hate-Speech-in-Southeast-Asia-New-Forms-Old-Rules.pdf>

narrative that Prabowo's son (Jokowi's opponent in 2019) is from the LGBT group, resulting in the labeling of supporters of same-sex marriage. This politicization of identity ultimately gave rise to hate speech and had an impact on the physical space where there was a lot of intimidation and threats of violence against LGBT groups.

In the Philippines, the incidence of hate speech targeting women is notably significant. Hatespeech in the form of laughing emojis were indicative of a mocking tone within conversations, with both sides (those who oppose misogyny and those who do not) ridiculing one another. The rampant use of vomit and angry emojis showcased the convoluted nature of the conversations. Top keywords showed that the majority of the conversation derided women: words like *"pokpok"* and *"puta"* were among the largest. Keywords like dress, modestly, and decent showed that a large part of discourse related to how women should dress. The word weak stood out as a stereotype falsely as-associated with women.<sup>52</sup>

During the campaign period leading to the 2016 national elections in the country, a woman who posted on social media that she will not vote for a lazy candidate got the ire of the candidate's supporters. For that post, she received hateful messages including rape and death threats. Maria Ressa, a journalist and CEO of the social media news site Rappler, has many times been the subject and target of hate speech by trolls and supporters of the President. "You are so ugly, you shouldn't have been born." "I hope you get raped." These are but some of the utterances she had been bombarded with for stating facts about the government and the president.<sup>53</sup>

These various identities ultimately become vulnerable. Hate speech maintains and re-establishes the target population's inferior status. In addition, hate speech occasionally encourages a negative view of others or a comparison between oneself and the other (Zamri et al., 2020). Based on race, ethnicity, religion, gender, we need to see how in social life, social constructions are formed which give rise to majority groups. Due to similarities in views, beliefs, characteristics and/or bodily conditions, the majority group ultimately has the power to determine.

This problem will continue to put pressure on groups with fewer numbers. In fact, this dangerous impact will not only affect certain groups. But it will also affect the stability of the country. Intimidation is closely linked to relations of influence and power. Parties who have stronger influence and power intimidate parties who have weaker influence and power or have no power at all. Intimidation is carried out by parties who have power against vulnerable and subordinate parties. In the context of fulfilling digital rights which should be universal, content moderation is not carried out proportionally and actually endangers minority groups.<sup>54</sup>

---

<sup>52</sup> BIG DATA ANALYSIS ON HATE SPEECH AND MISOGYNY IN FOUR COUNTRIES: BANGLADESH, INDONESIA, THE PHILIPPINES AND THAILAND  
<https://asiapacific.unwomen.org/sites/default/files/2023-03/ap-pve-big-data-misogyny-brief-s.pdf>

<sup>53</sup> Understanding Hate and Hate Speech: The Philippine Context  
<https://fma.ph/2022/06/29/understanding-hate-and-hate-speech-the-philippine-context/>

<sup>54</sup> perludem

Human rights, which are believed to have universal value, are actually the basic foundation for forming a regulatory framework for the protection of citizens in a democracy. Human rights have universal values which are then translated into various national legal products in various countries to protect and uphold human values. Not only majority groups, but also minority groups.

The Universal Declaration of Human Rights also noted the points of human rights protection including freedom and equality, the right to protection, and the right to participate in democracy. So, everyone has the right to a safe space in the democratic process. Article 1 affirms that 'All human beings are born free and equal in dignity and rights'. Rights that small groups of people believe ought to be recognised as human rights are sometimes referred to as 'moral rights'. They become recognised as human rights when there is a widespread consensus that they should be applied globally<sup>55</sup>

International initiatives have provided a growing body of standards and recommendations to guide government efforts to combat intolerance and hate speech. In particular, Human Rights Council Resolution 16/18 sets out a universally agreed action plan by states for addressing prejudice based on religion or belief.<sup>56</sup>

## Conclusion

Based on the comparative study above, we can draw some useful conclusions to better understand the regulatory framework in the region's democratic countries. First, hate speech against vulnerable groups is a common problem in the region. Social media has not been a safe space during political campaigns for racial and gender minorities (Indonesia & Malaysia), religious and gender minorities (Malaysia), as well as women who are political opposition or critical journalists (Philippines).

Second, there are differences in views between countries regarding the need for special regulations governing hate speech. Indonesia and Malaysia have quite strict regulatory instruments where hate speech is regulated in more than one legislative product. Meanwhile in the Philippines, only one legislative product prohibits hate speech. Election organizers in Indonesia and Malaysia also have special regulations prohibiting hate speech during elections, while the Philippines does not.

Third, the regulatory framework for hate speech that these countries have is still very limited, but excessive. This means that the groups protected from hate speech are limited to racial, ethnic and religious minority groups. In fact, hate speech also targets other vulnerable groups. On the other hand, the punishment and its application are very excessive. Both the Criminal Code and EIT Law in Indonesia, the Sedition Act in Malaysia, and the Cybercrime Prevention Act in the

---

<sup>55</sup> The Right to Safety: Some Conceptual Practical Issues [https://www.ucl.ac.uk/hazard-centre/sites/hazard\\_centre/files/wp9.pdf](https://www.ucl.ac.uk/hazard-centre/sites/hazard_centre/files/wp9.pdf)

<sup>56</sup> Malaysia: An inclusive policy measure is needed to end hate speech and discrimination <https://www.article19.org/resources/malaysia-inclusive-policy-measure-needed-to-end-hate-speech/> (23 September 2023)



Philippines have multiple interpretations and are prone to being misused to silence freedom of expression.

Therefore, this research recommends two things: domestic policy and foreign policy for the three countries. For domestic policy, the three countries can adopt a multistakeholder approach through the establishment of a social media council (SMC) as an independent state institution. SMC can become a multi-stakeholder forum to discuss internet content governance issues, filled with important actors, such as social media companies, CSOs that focus on human rights and democracy issues, researchers, digital technical experts, journalists and media, and of course groups that represent interests. vulnerable groups. SMC can encourage preventive measures and take action based on the authority of social media companies, as a replacement for the existing excessive penalties for spreaders of hate speech which are often used as a means of criminalizing expression.

For foreign policy, the author encourages the three countries to be more active in encouraging the presence of content governance policies based on human rights standards in ASEAN. These regional organizations are a potential force for pressuring multinational social media companies to do more to prevent the spread of hate speech. Of course, there are concerns that ASEAN's intervention in content governance could potentially threaten freedom of expression on the internet considering the decision-making process is based on consensus and the large number of authoritarian or semi-authoritarian countries in the region. Precisely because of this, the triumvirate of democracy champions in the region must be present first in maintaining a digital space that is safe and based on human rights. The "Maphilindo" spirit of cooperation that has been carried out to protect the Sulu Sea from the threat of extremism can also be transmitted to cooperation in protecting cyberspace from hate speech against vulnerable groups.

## References

- Aarons, Lauren, and Gabor Rona, 'State Responsibility to Respect, Protect and Fulfill Human Rights Obligations in Cyberspace', *Journal of National Security & Policy*, 8.2015 (2016), 503–30
- Amnesty International, 'Philippines: "Cybercrime" Law Threatens Free Speech and Must Be Reviewed', 2012 <<https://www.amnesty.org/en/latest/news/2012/10/philippines-cybercrime-law-threatens-freedom-expression-and-must-be-reviewed/>>
- Asia Centre, *Hate Speech in Southeast Asia: New Forms, Old Rules*, 2020
- Audrine, P, and I Setiawan, 'Dampak Regulasi Moderasi Konten Terhadap Kebebasan Berekspresi Di Indonesia' (Jakarta, 2021) <<https://repository.cips-indonesia.org/ms/publications/347662/dampak-regulasi-moderasi-konten-terhadap-kebebasan-berekspresi-di-indonesia>>
- Elumalai, Nalini (ARTICLE 19), 'Malaysia: An Inclusive Policy Measure Is Needed to End Hate Speech and Discrimination', *Article 19*, 2023 <<https://www.article19.org/resources/malaysia-inclusive-policy-measure-needed-to-end-hate-speech/>> [accessed 23 September 2023]
- Foundation for Media Alternatif, 'Understanding Hate and Hate Speech: The Philippine Context', *Foundation for Media Alternative*, 2022 <<https://fma.ph/2022/06/29/understanding-hate-and-hate-speech-the-philippine-context/>> [accessed 30 September 2023]
- Hamzah, Masjaliza, et. al, *Social Media Monitoring of Malaysia's 15th General Election* (Kuala Lumpur, 2023) <<https://www.facebook.com/CIJ.MY>>
- Sherly Haristya, *Moderasi Konten Dan Pemangku Kepentingan Lokal Di Indonesia*, 2022
- Roberts, Sarah T, 'Content Moderation', 2017 <<https://escholarship.org/uc/item/7371c1hf>>
- Robinson, Sally, and Jan Idle, *The Right to Safety: Some Conceptual Practical Issues*, *The Routledge International Handbook of Children's Rights and Disability*, 2003 <<https://doi.org/10.4324/9781003056737-48>>
- Roos Niza, Mohd Shariff, 'Sedition Act As a Pillar of Strength in Maintaining Unification for Multiracial Nation', *International Journal of Business, Economics and Law*, 6.4 (2015), 1–5 <[www.malaysia-today.net/the-sedition-act-1948/](http://www.malaysia-today.net/the-sedition-act-1948/)>

- Rosana, Ellya, 'Hukum Dan Perkembangan Masyarakat', *Jurnal Tapis*, 9.1 (2013), 104  
<<http://ejournal.radenintan.ac.id/index.php/TAPIS/article/view/1578/0>>
- Safitri, Ria, 'Undang-Undang Informasi Dan Transaksi Elektronik Bagi Perguruan Tinggi', *SALAM: Jurnal Sosial Dan Budaya Syar-I*, 5.3 (2018), 197–218 <<https://doi.org/10.15408/sjsbs.v5i3.10279>>
- UN Women , 'Big Data Analysis On Hate Speech And Mysogyny In Four Country: Bangladesh, Indonesia , The Philippines and Thailand' (UN Women, 2020), pp. 1–6
- Zamri, Norena Abdul Karim, Nur Nasliza Arina Mohamad Nasir, Mohammad Nurhafiz Hassim, and Syaza Marina Ramli, 'Digital Hate Speech and Othering: The Construction of Hate Speech from Malaysian Perspectives', *Cogent Arts and Humanities*, 10.1 (2023) <<https://doi.org/10.1080/23311983.2023.2229089>>