



First Edition | March 2024

Monitoring Report on 'Hate Speech', Information Manipulation, risks to freedom of expression and election integrity in the 2024 Election Period

Report Authors

Ratna Aini Hadi

Shinta Ressmy

Beltsazar Krisetya

Table of Contents

1.	Introduction	3
2.	Applicable freedom of expression standards	4
3.	Overview of Freedom of Expression in Indonesia	5
4.	Information manipulation/disinformation: Leading up to the 2024 Elections	7
5.	'Hate Speech' Monitoring	21
6.	Discrimination Against Human Rights Defenders (HRDs) and Civil Society Organizations	47
7.	Attacks Against Journalists	49
8.	The use and choice of social media platforms by presidential candidates	51
9.	Spending on Advertising on Meta.....	52
10.	Content and Account Takedown Requests	53

1. Introduction

The Damai coalition has created a brief monitoring report highlighting the current issues surrounding information manipulation, ‘hate speech’ and threats to freedom of expression and election integrity in the context of the 2024 elections.

More specifically we present the results of monitoring regarding (i) ‘misinformation’ and ‘disinformation’ or ‘information manipulation’ as monitored by MAFINDO; (ii) ‘hate speech’ as monitored by SAFEnet and the Alliance of Independent Journalists (AJI); (iii) discrimination against journalists as monitored by AJI; (iv) the use and choice of social media platforms by presidential candidates and election management bodies; (v) spending on political advertising on Meta; and (vi) government content and account takedown requests.

We observe that the human rights risks in the online environment and the proliferation of problematic content (such as disinformation, ‘hate speech’, or online harassment) are similar to those observed in the context of the 2019 elections. While those risks might be amplified in the online environment and social media companies ought to do more to live up to their responsibilities to respect human rights, we believe that the root causes of the problematic content observed in the online environment are related to political

polarisation and tensions in society, the marginalisation of certain groups and restrictions on freedom of expression and media freedom.

2. Applicable freedom of expression standards

Indonesia guarantees freedom of expression through its constitution and the Law on Human Rights (Law No. 39/1999). Indonesia is also a party to the International Covenant on Civil and Political Rights (ICCPR) which guarantees freedom of expression in its Article 19.

While we refer to the concepts of ‘hate speech’ and ‘disinformation’ in this report, they do not have any internationally agreed definition. Any restrictions on these types of speech must comply with the three-part test under Article 19(3) of the ICCPR.¹ The falsity of information is not *per se* a basis for restrictions unless it is connected to one of the legitimate aims listed in the three-part test, such as the protection of public health or national security.

Based on the data obtained by SAFEnet, several protected groups are targeted by ‘hate speech’, which can be understood in its broadest form as

¹ *The restriction must be (i) provided for by law, with sufficient precision to enable individuals to conduct themselves in adherence with the law; (ii) in pursuit of a legitimate aim, listed exhaustively as: (1) respect of the rights or reputations of others; (2) the protection of national security or public order (ordre public) ; (3) or of public health or morals; and (iii) Necessary in a democratic society, meaning that the State must specifically demonstrate*

any expression of hate towards an individual or group defined by a protected characteristic. For ‘hate speech’, Article 20(2) of the ICCPR requires that governments restrict speech that constitutes incitement to discrimination, hostility, violence or international crimes. The UN Rabat Plan of Action, adopted in 2012 establishes a six-part test for determining whether expression rises to the level of “incitement to hatred” which may be prohibited. The six-part threshold test considers (1) context of the issues, (2) status of the speaker, (3) intention of the speaker to incite violence against a target group, (4) content and form expression, (5) extent of its dissemination, and (6) likelihood of harm, including imminence of violence and discrimination.² In addition, the requirements of the three-part test under Article 19(3) of the ICCPR must always be fulfilled when limiting expression.

3. Overview of Freedom of Expression in Indonesia

Articles 23 of Law No. 39 of 1999 regarding Human Rights³ in Indonesia guarantees the freedom for each individual to select their political representatives and hold their own political beliefs. Article 25 of the same Law guarantees the right to publicly express one’s opinions. Despite these

² “One-pager on ‘incitement to hatred’”, United Nations Human Rights Office of the High Commissioner, https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/SeminarRabat/Rabat_threshold_test.pdf

³ (1999). Undang-undang Republik Indonesia Nomor 39 Tahun 1999 Tentang Hak Asasi Manusia. <https://www.komnasham.go.id/files/1475231474-uu-nomor-39-tahun-1999-tentang-H9FVDS.pdf>

guarantees, several pieces of domestic legislation are at odds with the right to freedom of expression.

The first such legislation is *UU ITE* or Law No. 11 of 2008 on Electronic Information and Transactions, which contains articles concerning defamation, ‘hate speech’, and false news⁴. The law has undergone two revisions, most recently in December of 2023. The law widely faced criticism due to several reasons. One being its broad and vaguely defined provisions, which critics argue can be applied in ways that infringe upon freedom of expression. The law’s enforcement has been criticized for being used as a tool to silence government critics rather than targeting genuinely harmful online behavior. Furthermore, the law’s harsh penalties for online defamation have been condemned as severe. From 2016-2020, SAFEnet’s data⁵ shows that cases under the ITE Law demonstrate a high conviction rate of 96.8% (744 out of 768 cases), with imprisonment reaching 88% (676 cases) resulting in prison sentences ranging from 1-5 years. The criticisms of the law have prompted a second revision of the legislation, which was recently passed by the DPR (People’s Representative Council) in December 2023, promising improvements to the shortcomings of the law.

⁴ (2023, December 4). Indonesia: Revise the Electronic Information and Transaction Law. *Article 19*. <https://www.article19.org/resources/indonesia-revise-the-electronic-information-and-transaction-law/>

⁵ Juniarto, D. (n.d.). Revisi UU ITE Total Sebagai Solusi. *SAFEnet*. <https://safenet.or.id/id/2021/03/revisi-uu-ite-total-sebagai-solusi/>

Apart from the ITE Law, freedom of expression in Indonesia has been limited in terms of platforms since the implementation of Minister of Communication and Information Regulation No. 5 of 2020 on Private Electronic System Operators. Through this regulation, the Indonesian government is granted powers to regulate Private ESO activity, gives authorities access to user data, and provides for sweeping notice and takedown orders⁶. This regulation is contrary to both Article 17 and 19 of the ICCPR as it lacks safeguards and can easily be misused by the State to interfere in citizens' private affairs and their expression rights in the online realm.

4. Information manipulation/disinformation: Leading up to the 2024 Elections

MAFINDO is an independent fact-checking organization that has been conducting verification of information regularly since 2016. The verified information originates from monitoring by fact-checkers and reports from the public. A team consisting of three coders verifies information using the quantitative content analysis method. Quality control is maintained through intercoder reliability to ensure the reliability and quality of the data. MAFINDO

⁶ (2021, September). Legal Analysis - Indonesia: Regulation of the Minister of Communication and Informatics Number 5 of 2020 on Private Electronic System Operators (Ministerial Regulation 5). *Article 19*. <https://www.article19.org/wp-content/uploads/2021/09/Legal-Analysis-Indonesia-Ministerial-Regulation-5.pdf>

refers to the 7 types of mis- and disinformation (Fig. 2) categorized by Wardle (2019)⁷ which denotes that the complexities of intention behind false information proposes that the categorization of mis- and disinformation resembles more of a spectrum rather than strictly binary. MAFINDO understands “misinformation” as false information that is created and spread without malicious intent and “disinformation” as false information that is created and spread with the intention of manipulating and misleading public perception, potentially to the extent of harming certain individuals/groups. For the purpose of this report, we focus on the frequency of all types of false mis/disinformation debunked by MAFINDO, regardless of where they fall in the spectrum of mis- and disinformation.

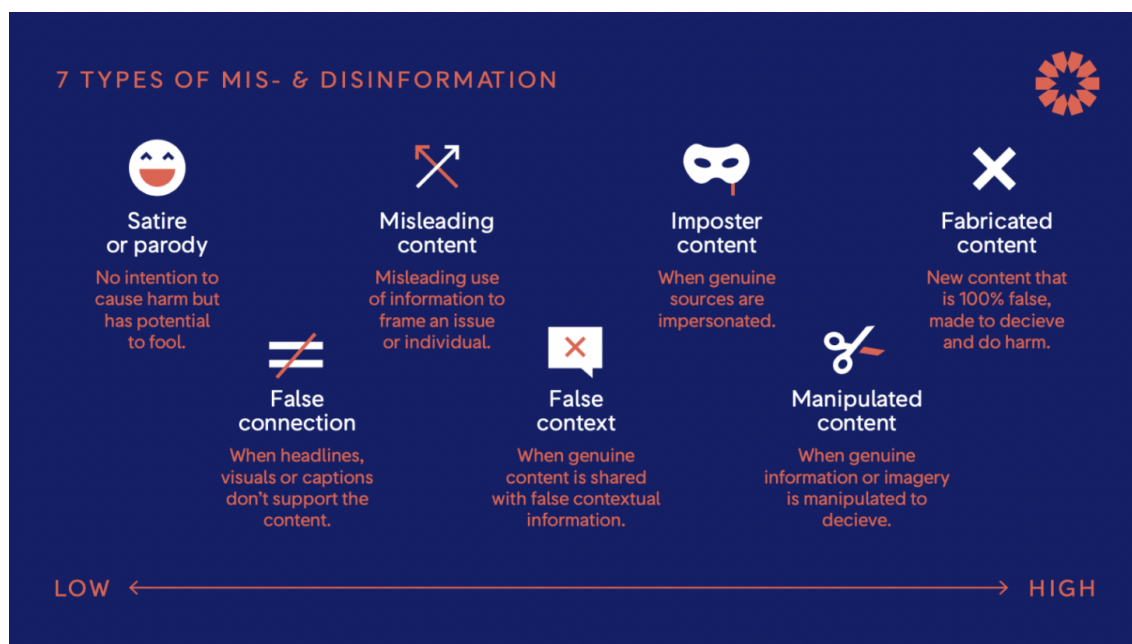


Figure 1. 7 Types of Mis- & Disinformation (Source: Wardle, 2019)

⁷ Wardle, C. (2019). Understanding Information Disorder. *First Draft*. 1-32.

In the previous presidential election season of 2019, MAFINDO reports⁸ that mis/disinformation appeared both before and after election day, in fact significantly increasing post-election day. The 2019 report highlights that the top four targeted groups by disinformation in 2019 are candidate supporters, candidates, KPU (the General Elections Commission), and the police.

In the year of 2023, a total of 2.330 mis- and disinformation were verified by MAFINDO. 55.5% of the identified disinformation (1292) are political-themed, and 27.7% (646) are election-themed (Fig. 2). Throughout 2023, an average of 53 election-themed disinformation was found and verified by MAFINDO. October to December 2023 especially experienced a rise in frequency of disinformation, especially in line with the start of the election process with the registration of presidential candidates in October and the start of the campaigning period from November to February 10th 2024, days before the election date. January 2024 sees the most disinformation in a month since January 2023 (155).

⁸ Purnama, F. Y., Murfianti, F., Hidayah, N., Wahid, A., Wijayanto, Kurniawan, H., Shaleh, M. H., Perangin-angin, L. L. K., Jannah, R. T., Arifin. S., Agustyati, K. N. (2023). Tipologi Hoaks Pilpres 2019. Masyarakat Antifitnah Indonesia.

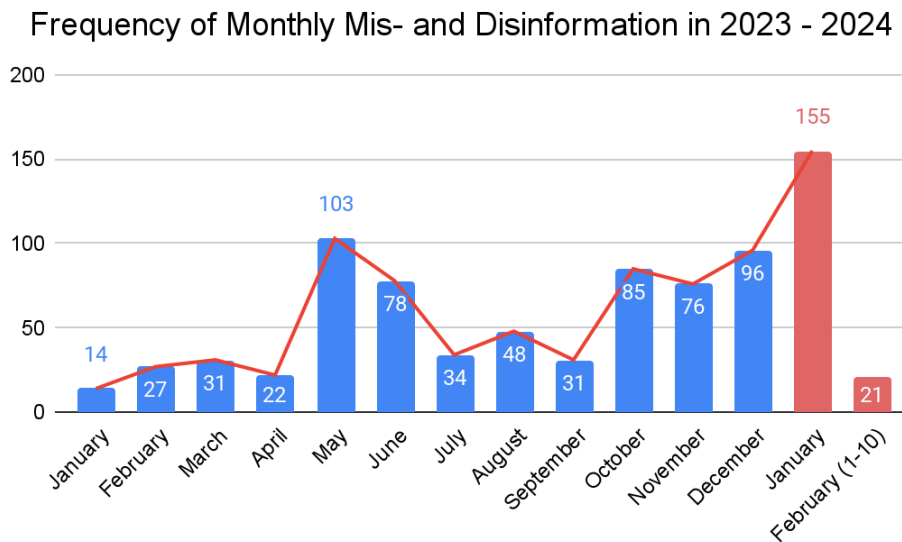


Figure 2. Frequency of monthly election-themed mis/disinformation (Source: MAFINDO)

For the purpose of this report, the period from 23 October 2023 to 10 February 2024 is monitored. This timeframe encompasses both the period before and during the campaign, which commenced on 28 November 2023. The most common type of false information found during this period is misleading (223 or 56%) and manipulated content (136 or 34.2%) (Fig. 3). The four top social media platforms where disinformation was found are (Fig. 4): YouTube (132), TikTok (97), followed by Facebook (87) and X/Twitter (40). This suggests that there is a growing trend of increasing audiovisual disinformation compared to solely text-based, making it more complex to fact-check.

Frequency of Mis- and Disinformation Based on Type of Content

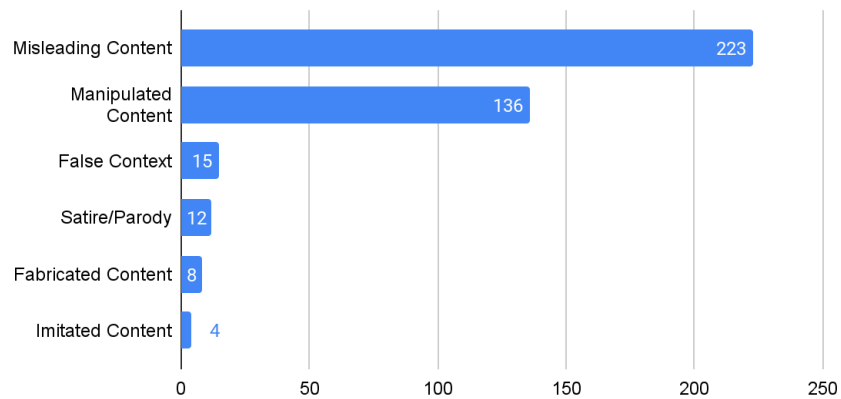


Figure 3. Frequency of disinformation based on type of content (Source: MAFINDO)

Frequency of Mis- and Disinformation Based on Social Media Platform

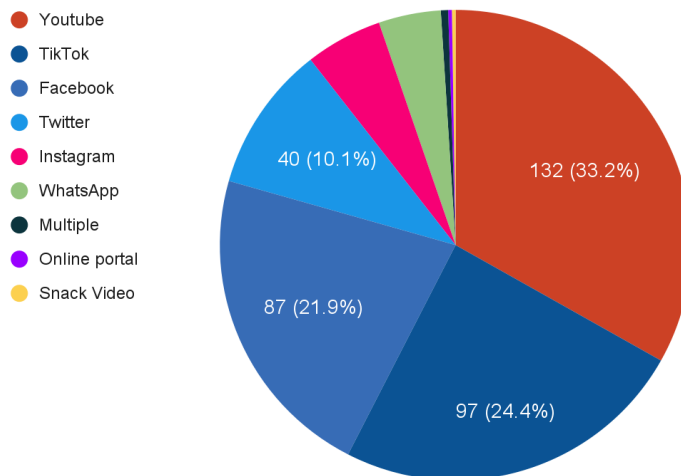


Figure 4. Frequency of disinformation based on social media platform (Source: MAFINDO)

Three disinformation strategies were identified in the period of monitoring; disinformation targeted 1) to delegitimize candidates, 2) to garner support for candidates, 3) the election process.

4. 1. Candidacy Delegitimization

All phases of the election process are targeted by disinformation. A tactic of information manipulation found in the beginning of the electoral process is a strategy of undermining the legitimacy of a candidate's candidacy.

Examples:

1. [Incident 1](#): In an attempt to delegitimize the candidacy of Gibran Rakabuming Raka, son of President Joko Widodo and running mate of presidential candidate from the Gerindra Party, Prabowo Subianto, there was ongoing discourse questioning the validity of Gibran's bachelor's degree. Some allegations suggest that his bachelor's degree diploma may be fake or equivalent to a vocational high school diploma instead. In November 2023, a video surfaced on [Facebook](#) and [Youtube](#) with the title "Gibran admits he is only a vocational high school graduate and apologized in front of the media". The video uses an image of Gibran in front of the media from an unrelated event and incorporates narration from [an article](#) quoting accusations from other parties.



Figure 5. Manipulated video of Gibran (Source: turnbackhoax.id)

2. [Incident 2](#): In November 2023, a user posted a TikTok video on the Facebook group called Indonesia Bersuara, about presidential candidate, Anies Baswedan. [The post](#) targeted Anies' personal identity, claiming that he and D.N. Aidit (a scrutinized Indonesian communist politician) are both descendants of Yemen and have the ambition to dominate the Republic of Indonesia. This narrative attempts to create a connection between Anies and communism.

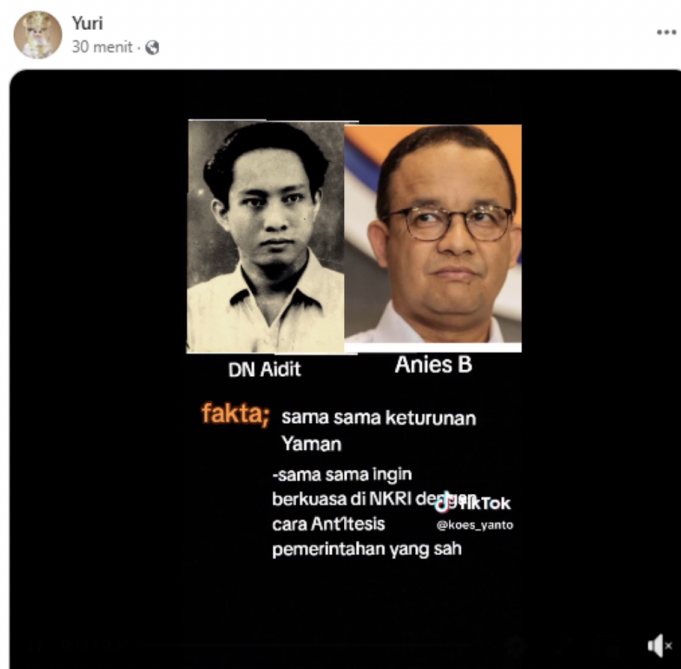


Figure 6. Disinformation regarding connection between Anies and D.N. Aidit (Source: turnbackhoax.id)

A few manipulated pictures were produced to delegitimize the candidacy of Ganjar Pranowo, the presidential candidate from the Indonesian Democratic Party of Struggle.

Examples:

3. [Incident 3](#): A manipulated picture of Ganjar as a cover model of Playboy magazine was posted on Facebook (Fig. 7). The manipulated picture was originally from the November 1976 issue of Playboy Magazine.

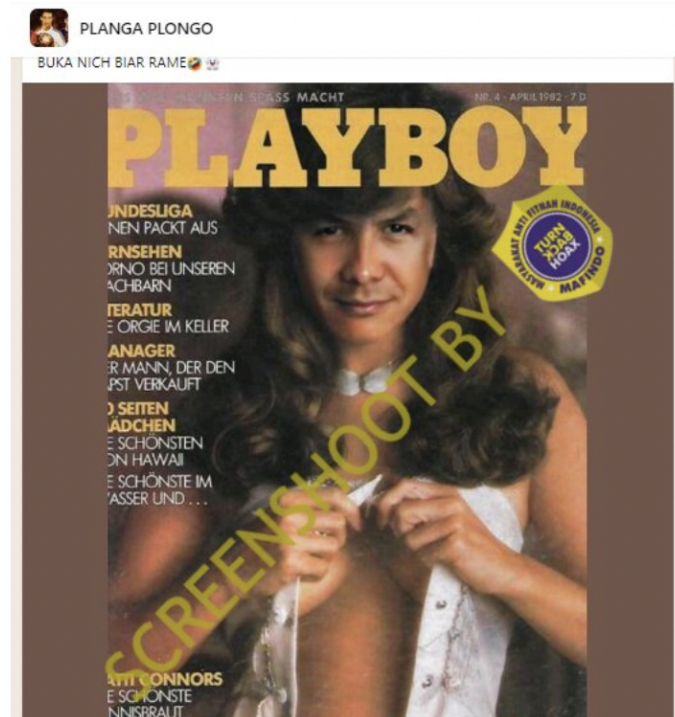


Figure 7. Manipulated photo of Ganjar as playboy cover (Source: turnbackhoax.id)

4. [Incident 4](#): A picture of Ganjar and Mahfud MD, his running mate, taken during the mandatory health check-up for presidential and vice-presidential candidacy was altered to create the false impression that they were at a psychiatric hospital instead.

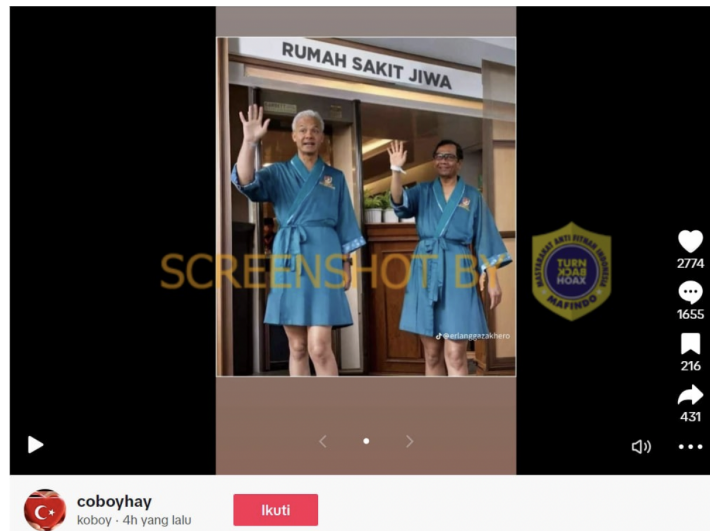


Figure 8. Manipulated photo of Ganjar and Mahfud at psychiatric hospital (Source: turnbackhoax.id)

These examples of false and manipulated information are aimed at weakening public support to the candidates through sowing doubt and disbelief.

4. 2. Enticing Support for Candidates

Another trend of disinformation is aimed at garnering support for certain candidates by spreading false information regarding the support of key opinion leaders such as public figures, government officials, and celebrities towards certain presidential candidates. These types of disinformation are used to legitimize the candidacy of the candidates, shape public opinion to garner support for said candidates.

Examples:

1. [Incident 1](#): In January, a user posted a picture on Facebook claiming that Indonesia's Minister of Finance, Sri Mulyani, is endorsing the pair Ganjar-Mahfud in the elections. Previously, the fact-checking team had verified prior disinformation regarding Sri Mulyani's alleged support for Ganjar.



Figure 9. Disinformation regarding support for Ganjar (Source: turnbackhoax.id)

2. [Incident 2](#): In November, a TikTok user posted a video claiming that the Chairman of the High Council of the Democratic Party, former president

Susilo Bambang Yudhoyono, had granted approval for party volunteers to support Ganjar in the 2024 election.



Figure 10. Disinformation regarding support for Ganjar (Source: turnbackhoax.id)

3. [Incident 3](#): A TikTok user posted a video claiming that the famous 37-year-old, Indonesian actor, Raffi Ahmad is a supporter of candidate Anies and his running mate, Muhaimin. This is one example of a series of posts on social media found alleging support from celebrities to different presidential candidates.



Figure 11. Disinformation regarding support for Anies (Source: turnbackhoax.id)

4. 3. Targeting the Electoral Process

A trend of disinformation found in the previous 2019 elections as well is attacking the electoral process. A number of disinformation was found targeting the Election Management Bodies, trying to undermine the electoral process.

Examples:

1. [Incident 1](#): The electoral process in Indonesia operates as follows; voters are required to perforate paper ballots to indicate their chosen candidates. [A post](#) on X (formerly known as Twitter) has surfaced, purportedly warning the public about the potential manipulation of voting ballots using sharp rings. Previously, a similar narrative regarding rings used to tamper with ballots were found in 2012 and 2019.



Figure 12. Disinformation regarding rings used to tamper voting ballots (Source: turnbackhoax.id)

Moving forward throughout the post election period, the public should anticipate and be aware of possible mis- and disinformation targeting all phases of the electoral process, such as false narratives regarding the casting of ballots and ballot counting. We will highlight more in the post-election monitoring briefing.

5. 'Hate Speech' Monitoring

The national motto of Indonesia is *Bhinneka Tunggal Ika*, “Unity in Diversity”. At its core, Indonesia is both a highly culturally and geographically diverse country with 1,340 recognized ethnic groups across over 18,000 islands. The extensive diversity, whether arising organically or through external factors, results in the emergence of minority communities. Minoritizing groups globally encounter diverse threats rooted in socio-political conditions. These challenges, influenced by factors such as polarisation, governance, societal structures, and political climates, contribute to the vulnerabilities faced by minority communities across different regions.

In their report regarding the government’s efforts to safeguard the rights of minority groups in Indonesia⁹, The National Human Rights Commission (2016)

⁹ Anam, C., et al. (2016). Upaya Negara Menjamin Hak-Hak Kelompok Minoritas di Indonesia (Sebuah Laporan Awal). *Komisi Nasional Hak Asasi Manusia*.

highlights the existence of minority groups based on: race, ethnicity, religion, people with disabilities, sexual orientation, and gender identity. Other minority groups include those based on gender and political ideology.

In order to safeguard unity in diversity in Indonesia, the term “SARA” (Suku, Agama, Ras, Antargolongan) is widely used in the context of politics and social discussions to discourage the exploitation of differences in ethnicity, religion, race, or group affiliations for negative purposes.

5. 1. Monitoring Dashboard

Working together with Monash Data & Democracy Research Hub and Boston University, AJI, the Alliance of Independent Journalists, have created a [‘hate speech’ monitoring dashboard](#)¹⁰ for the 2024 Indonesian General Election period. The monitoring dashboard divides ‘hate speech’ into five categories; 1) contains threat/incitement to violence, 2) contains insult, 3) contains identity attack, 4) contains sexually explicit words or phrases, and 5) contains obscenity or profanity. At 89.4% accuracy, this dashboard uses an NLP (Natural Language Processing) model to detect ‘hate speech’ on various social media platforms, namely; Twitter/X, Facebook, Instagram, and online

¹⁰ The dashboard was created as follows; 1) firstly, FGDs were held with possible vulnerable groups to gain further insight about their struggles, 2) data mining is done by scraping various social media platforms using targeted groups as keywords, 3) Data is manually annotated as base indicators to identify hate speech as training data, 4) Several NLP models are trained to create hate speech detection model, 4) data is visualized as dashboard. For more information, see Aliansi Jurnalis Independen. *2024 Indonesian General Election Hate Speech Dashboard*. Available at: [<https://aji.or.id/hate-speech-monitoring>]

articles. The ‘hate speech’ detected is directed towards potentially protected groups in Indonesia based on religion (Syiah, Ahmadiyah, Christian, Catholics), race (Chinese-Indonesian/Tionghoa and Rohingya), disability, as well as sexual orientation and gender identity (LGBTQ). Given the current underlying political context of the war in Gaza, the search also included the search terms Israelis/Jews.

In the period of 23 October 2023 to 31 January 2024, 26.86% of total data scraped, or 155,000 posts were detected to contain ‘hate speech’ (Fig. 13).

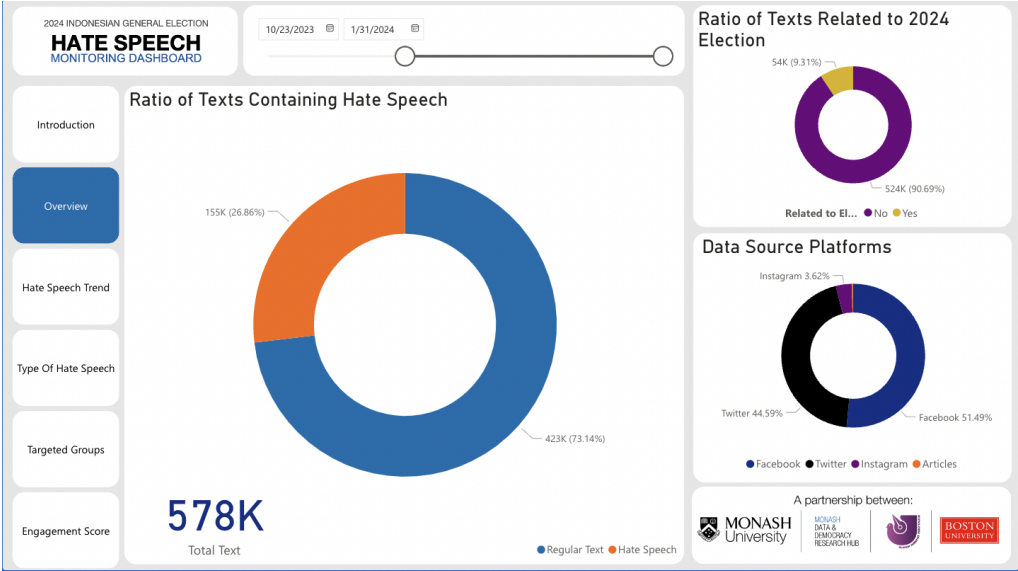


Figure 13. Monitoring dashboard: ratio of texts containing ‘hate speech’ (Source: AJI)

55,000 (51.5%) are found on Facebook, 96,000 (44.6%) are found on Twitter/X, and 4,000 (3.6%) are found on Instagram (Fig. 14). The dashboard

detected an overall spike in ‘hate speech’ found on Facebook, Instagram, X/Twitter and articles in December 2023.

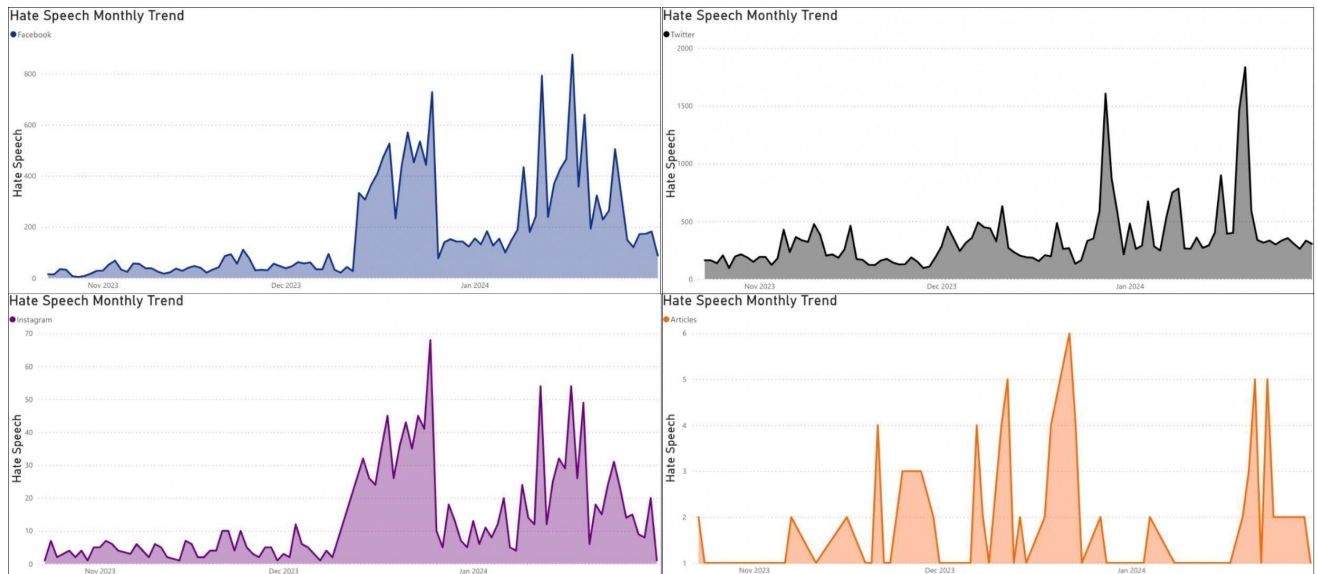


Figure 14. ‘Hate speech’ monthly trend based on social media platforms (Source: AJI)

The most common type of ‘hate speech’ identified contained insults (20,900), obscenity or profanity (3,600), and threats or incitement to violence (3,000). In the period monitored, the most targeted group is those with disability (13,100), Jewish (4,500), LGBT (1,400), and Chinese-Indonesian (1,300) (Fig. 15).

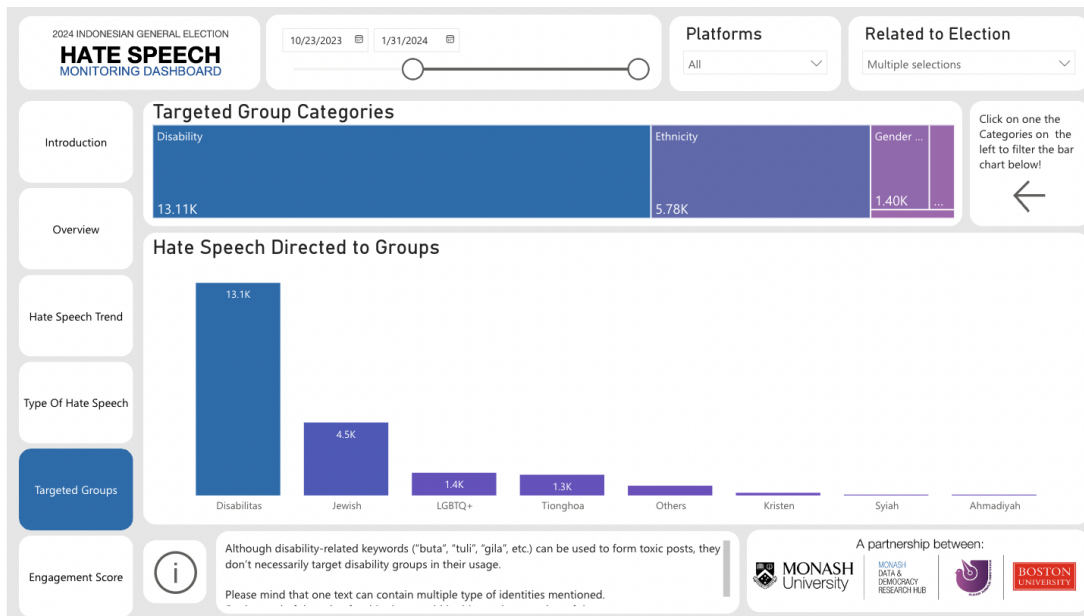


Figure 15. Frequency of ‘hate speech’ based on targeted groups (Source: AJI)

Figure 16 showcases how much reach on social media the ‘hate speech’ detected has; 17,000 tweets on X/Twitter gained 1 million likes and 425,000 retweets, 7,552 posts on Facebook gained 436 thousand shares and 2 million comments, 545 Instagram posts gained 16 million likes and 801,000 comments.

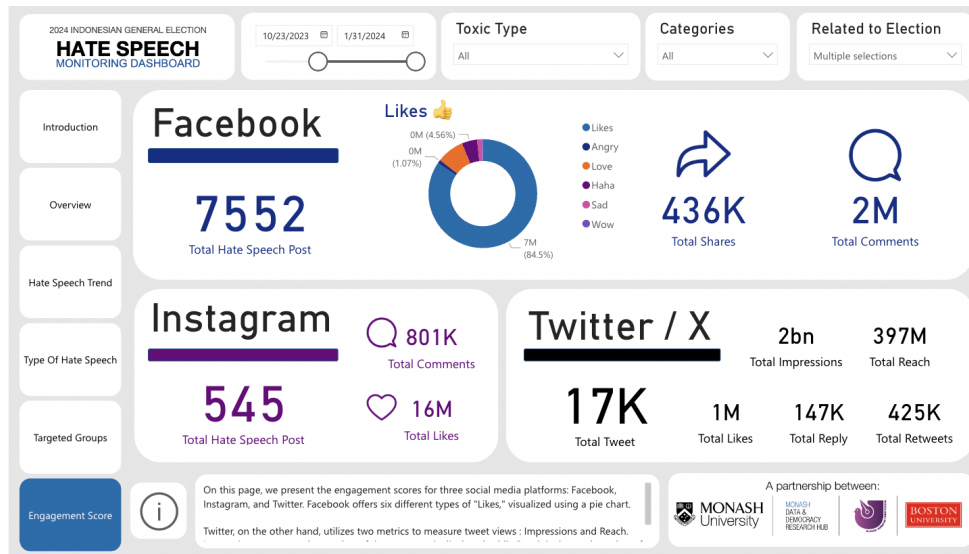


Figure 16. Trend of social media reach of hate speech (Source: AJI)

The dashboard highlights that ‘hate speech’ targeting protected groups on social media (X/Twitter, Facebook, and Instagram) is rampant during the election season. The dashboard concludes that insults are the most common type of ‘hate speech’ and the disabled are the most targeted protected group. ‘Hate speech’ trends tend to follow the trend of ongoing public discourse. In the discourse surrounding voting rights during the 2024 elections, an issue arises regarding the voting rights of persons with disabilities. Misinformation regarding the voting rights of the disabled community was widespread causing instances of ‘hate speech’ prevalent among those who disagree that people with mental disorders have the right to vote; further stigmatizing people with mental disorders. In reality, in 2019 the Constitutional Court has decided that people with mental disorders can exercise their right to vote as

long as there is no permanent mental or memory impairment and that the right to vote for people with mild to severe mental disorders are waived.

In the following section of data analysis regarding ‘hate speech’, we will explore how the trends of targeted ‘hate speech’ found on the dashboard relate to the ‘hate speech’ database that has been compiled by SAFEnet.

Many posts were at Level 1, Level 2 and Level 3 of the severity test. Comments that were inciting violence or hostile were on the low side. Negative portrayals of those targeted protected groups have been consistent over the years, including in the run-up to the elections. Monitoring has found that such conversations on social media lead to heightened hate speech, including instances of doxing and incitement of violence.

5. 2. ‘Hate speech’ Database

SAFEnet collects social media data on hate speech against protected groups during the election season. The data collection process involved both manual methods and the employment of Sonar (application) adhering to the ‘hate speech’ coding guidelines established by SAFEnet. SAFEnet refers to the United Nations Strategy and Plan of Action on Hate Speech (2019) to define ‘hate speech’, while acknowledging that this does not constitute a legal definition under international human rights standards. The typology and

categorization of hate speech were derived from the prescribed coding book, as provided in the attached document.

The data collection period starts from 23 November 2023 to 10 February 2024. The classification of ‘hate speech’ contains 3 conditions: 1) it targets the identity of a certain group, 2) it contains hostile legitimacy, and/or 3) the speech contains negative stigmatization of vulnerable groups (SAFEnet coding book, 2023).

For the purpose of this monitoring, we refer to the [Centre for Independent Journalism’s severity scale](#) (Fig. 17), which categorizes the levels of potential harm/severity of ‘hate speech’ as follows:

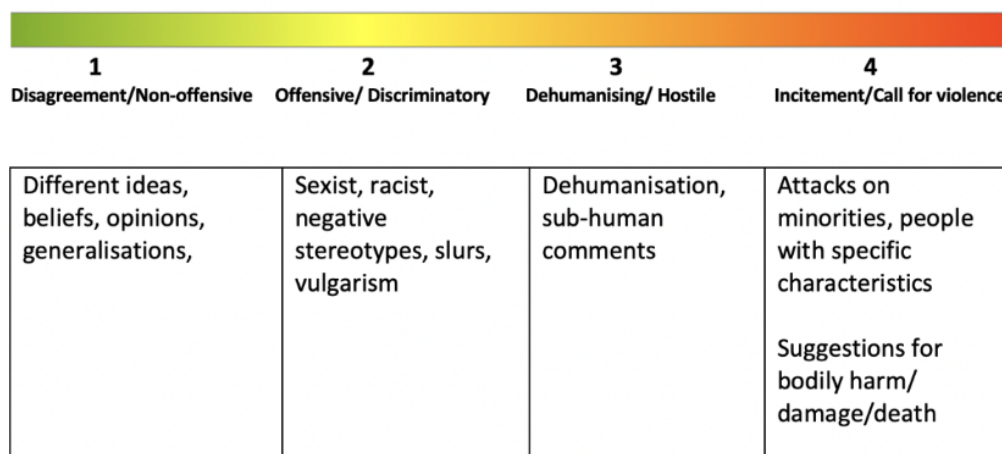


Figure 17. Parameters and severity scale of hate speech (Source: Centre for Independent Journalism Malaysia)

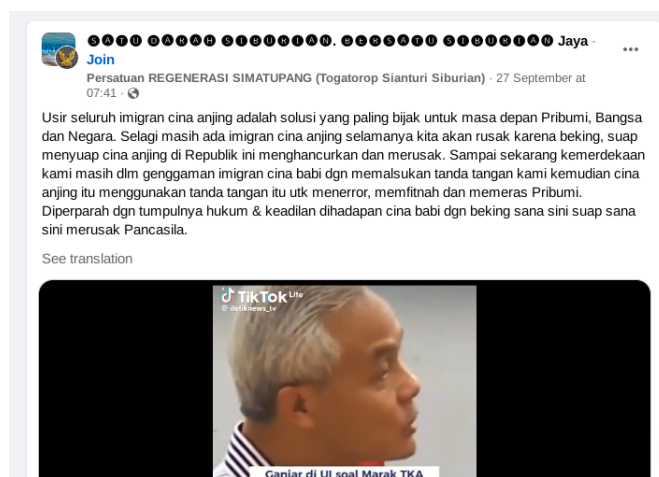
Based on the conditions of ‘hate speech’ categorization by SAFENet referring to the UN strategy, the collected data is in line with the parameters of ‘hate speech’ set by the CJI and can be categorized between the 2-4 levels of severity (Fig. 17). ‘Hate speech’ that is found ranges from being offensive, dehumanizing, to calling for violence.

The following are trends of ‘hate speech’ towards protected groups in Indonesia. We describe them under the different protected characteristics.

5. 2. 1. Race

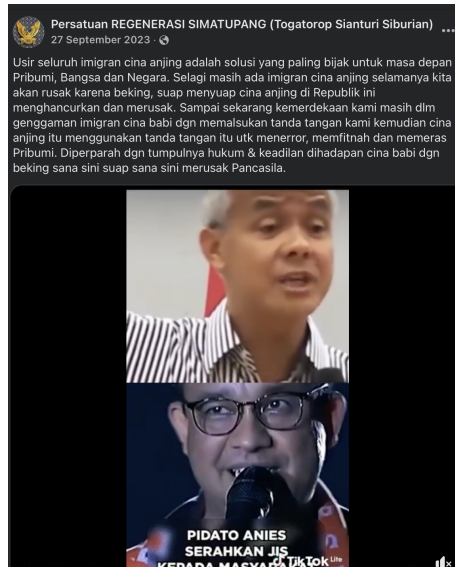
A common target of racist sentiments in Indonesia is the ethnic-Chinese community. The history of racism and negative rhetoric against the ethnic Chinese goes back to the scapegoating of Chinese Indonesians during the 1965 anti-communist purge and the 1998 Asian Financial Crisis.

1. The narrative that a certain political candidate is of Chinese-descent is used as a common tactic of misinformation and ‘hate speech’.

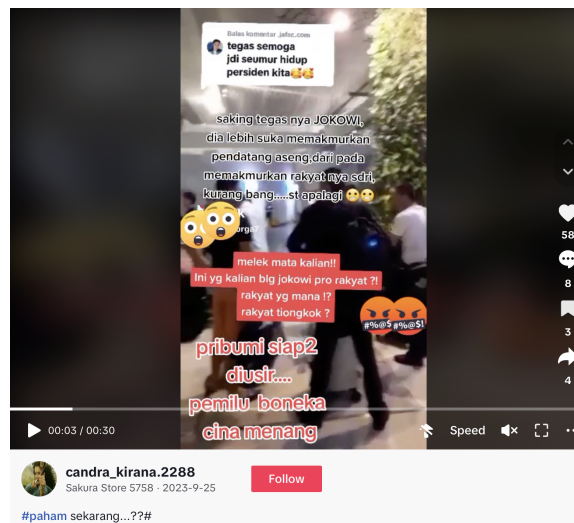


- a. The initial instance manifests within a November 2023 Facebook post wherein a group juxtaposes the discourses of Ganjar Pranowo and Anies Baswedan concerning the influx of foreign labor into Indonesia. This post (screenshot from TikTok's video) comprises content featuring 'hate speech', notably advocating for the expulsion of Chinese nationals and immigrants within Indonesia. Furthermore, it propagates allegations of potential blackmail by Chinese individuals upon the indigenous populace. Such expressions pose a substantial threat to the welfare and security of immigrants and foreign residents within Indonesia.

Harm test: Level 4, referring to the call to expel all Chinese citizens and people of Chinese descent. Apart from that, there were also insults and accusations that China had bribed Indonesia.



b. In the second example, a post was found uploaded on TikTok regarding accusations against Joko Widodo of prioritizing the interests of foreign immigrants. The immigrants referred to in the video are specifically ethnic Chinese. There are also narratives that fear that natives will be expelled from Indonesia and that the election victory will belong to China.



Harm test: level 2, referring to Joko Widodo's accusation of making foreigners more prosperous than natives and accusations that natives will be expelled from Indonesia.

2. Another narrative found is that certain political actors are supported by the Chinese, alluding to being pro-communism.



Another example of ‘hate speech’ was found in a candidate for DKI Jakarta legislative member, Chong Sung Kim, who is an Indonesian citizen of Korean descent. A TikTok account spreads fake news and ‘hate speech’ related to his candidacy. The account accused Joko Widodo of carrying out a massive exodus of the Korean-Chinese population to Indonesia. This brings a big narrative that behind Jokowi, there are the interests of foreign countries, in this case Korea and China. He also accused Chong Sung Kim of only changing citizenship a year ago, even though in fact Chong Sung Kim has been an Indonesian

citizen since 1992. This evidence underscores the presence of genuine racial discrimination within the electoral procedures in Indonesia. The proliferation of hateful narratives and dissemination of fake news possess the capability to undermine the standing of candidates and consequently foster negative societal perceptions.

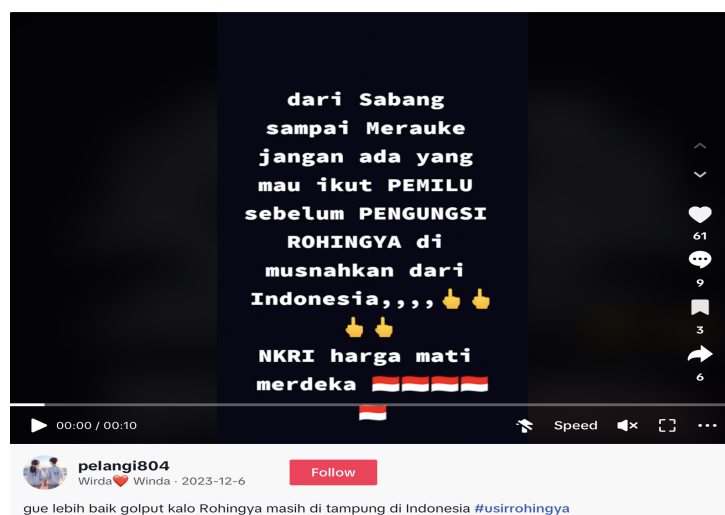
Harm test: level 2, referring to Joko Widodo's accusation of carrying out a mass exodus of Chinese-Korean citizens to Indonesia.

In 2017, a significant influx of Rohingya refugees, estimated at one million individuals, occurred, with many seeking asylum in multiple countries, including Malaysia, Bangladesh, Indonesia, and Australia. This migration was precipitated by pervasive ethnic discrimination and systematic human rights violations perpetrated by the government of Myanmar, which refused to recognize the Rohingya ethnic group and subsequently revoked their citizenship. Consequently, Indonesia's reception of Rohingya refugees in 2023 does not mark the inaugural instance of such migration to the nation.

As previously discussed, trends of 'hate speech' tend to follow the public discourse. Disinformation and 'hate speech' targeting Rohingya refugees were circulating around and during the electoral period¹¹. Conversations

¹¹ (2024, January 2). 'Rohingya di Sidoarjo', 'Rohingya minta tanah', 'Menlu Retno usir Rohingya' – Bagaimana narasi kebencian dan hoaks bekerja menyudutkan etnis Rohingya?. *BBC News Indonesia*. <https://www.bbc.com/indonesia/articles/c03y7n3k12lo>

about the opposition of and desire to expel Rohingya refugees from Aceh began to surface through narratives that demonize the refugees, including through tactics of disinformation and ‘hate speech’.



One instance of ‘hate speech’ manifests extensively on the TikTok platform, wherein there are calls for the expulsion of Rohingya refugees. The exemplified case depicts a threat to boycott elections should Rohingya refugees remain in Indonesia. This call is accompanied by the hashtag "#expelRohingya".

Certainly, Rohingya refugees, as a vulnerable demographic, are subject to discrimination and ‘hate speech’, particularly within the Indonesian context, exacerbated by the electoral climate. This situation is further compounded by the hostility expressed by certain segments of the population, who refuse to participate in the electoral process if the refugees remain within Indonesian territory.

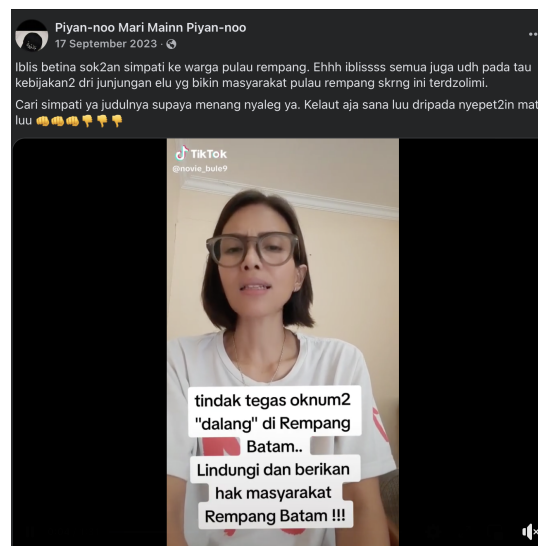
Harm test: level 4, referring to calls to exterminate Rohingya in Indonesia and the threat of an election boycott.

5. 2. 2. Gender

This type of content contains demeaning ‘hate speech’ against women. The content found during this period of monitoring is an attack against women involved in politics, both directly and indirectly, with the aim of undermining their integrity, dignity, self-worth, and status. This content contains accusations against these women, labeling them as mistresses.

1. This content contains accusations against these women, labeling them as mistresses. An example is an account that attacked one of the prospective legislative candidates who was taking a photo with Rempang residents and used a narrative of solidarity with Rempang residents after experiencing conflict. This photo then received a lot of negative comments because it was thought that it was only using Rempang residents to campaign and ask for support. Besides being called Mistress, the prospective candidate is also called a "female affiliated with the PKI [Indonesian Communist Party]" as well as various narratives that doubt her credibility if she runs for legislative office. In the uploaded content, ‘hate speech’ has been directed toward a female legislative candidate who was producing a video pertaining to the

conflict in Rempang. The post conveys the narrative that the aforementioned “betina” candidate is soliciting votes by exploiting the situation surrounding the Rempang conflict.



2. In another case, some of the ‘hate speech’ found against women, the word used to describe a woman is not *wanita*, but instead *betina*. *Betina* is commonly used to refer to animals or objects, not humans. This choice of language targets to diminish the status of women to the point where they are regarded as animals.

In the uploaded image, it was asserted that the depicted individual is a female affiliated with the Ummat Party. Specifically, her attributed qualities are delineated as being confined to those befitting a Legislative Member Candidate hailing from Jember, East Java. Additionally, she is purported to stand as a candidate associated with

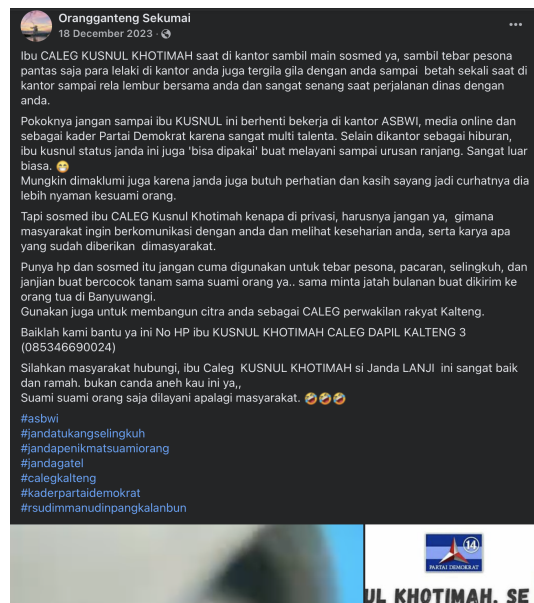
Sengkuni. It is noteworthy to recognize that Sengkuni occupies the role of the antagonist within Javanese wayang narratives, thereby imbuing the aforementioned candidacy with implicit symbolic connotations.



Harm test: level 3, referring to the narrative of "devil" and "*betina*" attached to the woman in both posts. The term female is usually applied to objects other than humans, for example, animals.

3. The image depicted portrays an instance where a legislative candidate in Kalimantan was accused of being a "widow who stole someone's husband." This accusation underscores the persisting social stigma against widows and reveals the prevalent discrimination against women

within the societal framework, particularly evident during electoral processes. Furthermore, the dissemination of the victim's personal photograph without her consent, accompanied by hashtags such as #jandapenikmatsuamiorang, #jandasukaselingkuh, and #jandagatel, serves to perpetuate accusations suggesting her inclination toward engaging with married individuals. Within the political context, such actions undoubtedly yield adverse implications for women's participation in democratic processes, fostering a climate of apprehension wherein women may hesitate to actively engage due to heightened susceptibility to 'hate speech' and discriminatory treatment.



Harm test: level 3, the discourse revolves around the dissemination of defamatory narratives portraying an individual as an intruder into the marital union of another, coupled with the unauthorized distribution of videos and photographs belonging to the rightful owner.

5. 2. 3. Disability

This type of content contains ‘hate speech’ against individuals with disabilities. Misinformation regarding the eligibility of people with mental disorders to vote in the elections was widely spread and fueled the eruption of ‘hate speech’ against people with mental disorders, furthering stigmatization. Outrage ensues claiming that “crazy people” have the right to vote in the upcoming elections. However, Article 148 of Law Number 36 of 2009 concerning health states that people with mental disorders (ODGJ) have the same rights as other citizens¹². During the 2019 Presidential Elections, the Constitutional Court emphasized that as long as there is no permanent mental or memory impairment, individuals with mental disorders can still exercise their voting rights. However, if the disorder is mild or severe, their voting rights are waived.

¹² N.d. (n.d.). CEK FAKTA: Hoaks KPU Mendata Orang Gila Ikuti Pilpres 2024. *Merdeka.com*. <https://www.merdeka.com/peristiwa/cek-fakta-hoaks-kpu-mendata-orang-gila-ikuti-pilpres-2024-38181-mvk.html?screen=9>



One example of ‘hate speech’ and discrimination against people with mental disorders is a netizen's post regarding the voting rights of people with mental disorders in the 2019 election. In this post, he argued that as a result of people with mental disorders being able to take part in the 2019 election, his madness has carried over to this day. In fact, the voting rights of people with mental disorders are guaranteed by the constitution as long as they are accompanied by a recommendation from a psychiatrist and if their mental disorders are still at a certain stage.

The utilization of the term "crazy people" is highly discriminatory and constitutes a manifestation of ‘hate speech’ directed toward marginalized communities. In lieu of employing the phrase "crazy people," it is imperative to adopt terminology such as "individuals with mental disabilities" or "individuals with mental disorders" in order to foster inclusivity and uphold respect for diverse populations.

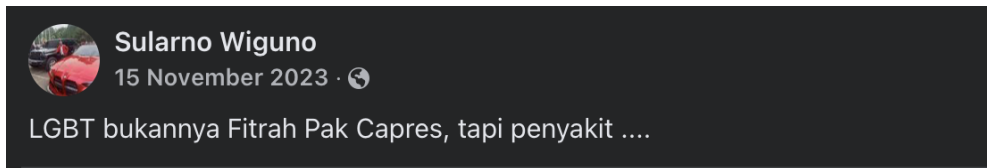
Harm test: level 2, the discourse pertains to the narrative of stigmatization against individuals with mental health conditions, who have been depicted as instigators of chaos within the electoral processes during the preceding two periods.

5. 2. 4. Sexual orientation, gender identity and intersex status

This content refers to ‘hate speech’ targeting the LGBT community. It includes the rejection of the existence of LGBT groups and any efforts to normalize LGBT activities. Instances of ‘hate speech’ against the LGBT community have been found on several platforms such as Facebook. In these posts, it is observed that there is an attempt to manipulate public opinion through allegations of the LGBT movement as a foreign threat that would endanger Indonesia.

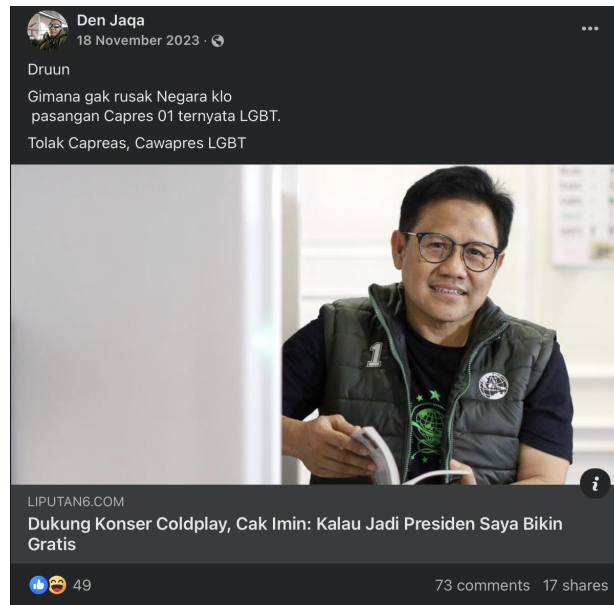
Examples:

1. In November 2023, a netizen corrected a statement by one of the presidential candidates regarding the anti-discrimination treatment of LGBT. The netizen uploaded a statement on his Facebook account with the narrative that LGBT is not a natural thing, but a disease. This indicates that there is discrimination and ‘hate speech’ towards LGBT as one of the protected groups in Indonesia.



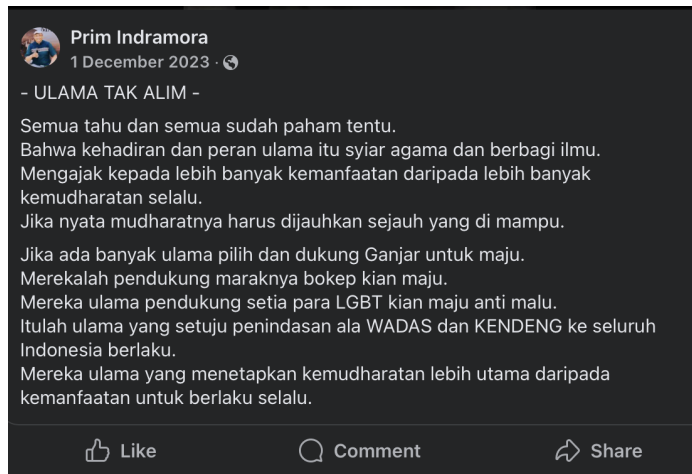
Harm test: level 2, because they think that LGBT is a disease, it is a stigma that is widely circulated in society, on the other hand, LGBT is trying to get rid of this stigma.

2. In light of Coldplay's scheduled concert in Indonesia coinciding with a statement made by Muhaimin Iskandar, the vice presidential candidate, proposing free admission to the event if elected, a polarized discourse emerged. One social media user levied an accusation against Muhaimin Iskandar, attributing alignment with LGBT ideologies to him due to Coldplay's known advocacy for LGBT rights. A Facebook post asserted that the endorsement of an LGBT-supportive vice presidential candidate could detrimentally affect the nation. It urged the populace to reject both the candidacy and the underlying LGBT advocacy. Certainly, Muhaimin's involvement in endorsing the Coldplay concert, including his attendance, consequently diminished his electoral prospects among homophobic and conservative factions.



Harm test: level 2, because there were LGBT accusations against Muhaimin Iskandar without any clear accusations.

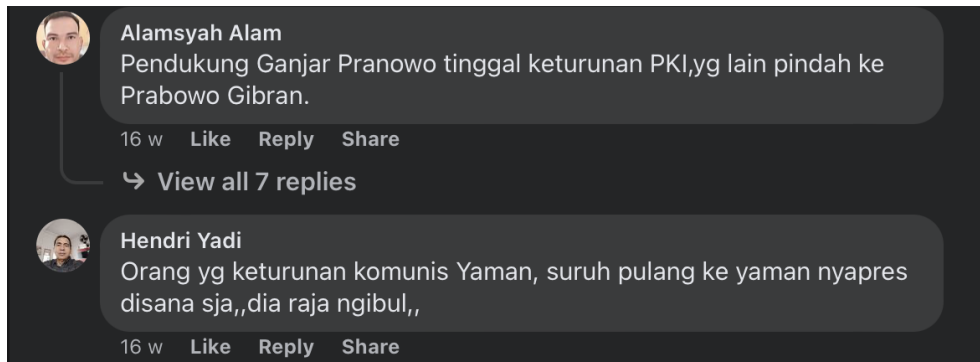
3. In December 2023, a netizen uploaded a status on Facebook containing 'hate speech' against one of the presidential candidates, Ganjar Pranowo. In the upload, it was also stated that the ulama (religious leader) who supported Ganjar Pranowo was the same as openly supporting LGBT and supporting the spread of pornography. It was also stated that these ulama agreed with oppression like what happened in Kendeng and Wadas to apply throughout Indonesia.



Harm test: level 2, contains slander against Ganjar Pranowo who is accused of perpetuating and supporting LGBT along with religious leaders.

5. 2. 5. Political opinion

Anti-communist sentiments in Indonesia emerged during the mid-20th century, intensified by events like the failed Madiun uprising in 1948 and the violent anti-communist purge in 1965. This period led to a lasting aversion to communism and restrictions on leftist activities. A common misinformation and 'hate speech' rhetoric during political elections is the allegations that a certain candidate is a communist and is affiliated with the Communist Party.



In November, netizen comments were found on Facebook stating and labeling Ganjar Pranowo's supporters as descendants of the PKI. There were also comments that responded containing 'hate speech', accusing Ganjar Pranowo of being a communist who came from Yemen and was expelled from running for president only in Yemen.

Harm test: level 2, because of the accusation that Ganjar's supporters are descendants of the Indonesian Communist Party and this has become a stigma in society.

5. 2. 6. Religious Minorities

'Hate speech' against religious minorities in Indonesia during the election period refers to acts of discrimination against certain groups such as Shia, Ahmadiyya, or other religions. sentiment towards religious minorities refers to various minority religious sects that are suspected of controlling the government. The historical roots of this discrimination start from Law Number 1/PNPS/1965, the Joint Ministerial Decree concerning the Establishment of

Houses of Worship, the classification of official and non-official religions to the heretical label from the Indonesian Ulama Council (MUI).



For instance, the uploaded content levies allegations suggesting that supporters of Anies Baswedan are affiliated with radical factions such as the FPI, HTI, Wahabi, and Shia, purportedly stemming from supporters of Prabowo Subianto. However, it is imperative to note that these accusations lack substantial evidence, thereby possessing the potential to engender societal discord and fragmentation.

Harm test: level 2, it is based on false information related to Anies Baswedan supporters originating from FPI, HTI, Wahabi, and Shia.

6. Discrimination Against Human Rights Defenders (HRDs) and Civil Society Organizations

In this section, our methodology entails the utilization of the SAFEnet dataset obtained from the SAFEnet complaint channel. The data collection spanned from October 13, 2023, to February 10, 2024. This dataset is deemed internal and has been disclosed solely for specified purposes, including its application within this research endeavor.

The phenomenon of online attacks against activists constitutes a grave infringement upon human rights, fundamentally imperiling the sacrosanct prerogative of individuals to freely articulate their opinions and expressions, whether through direct or indirect means. Egregiously, the prevalence of digital onslaughts persists unabated, particularly accentuated during electoral cycles, owing to a confluence of factors.

In the period leading up to the general election, instances of aggression targeting activists have been observed across various platforms. These instances encompass attacks conducted both by anonymous actors and directly perpetrated by the platforms themselves, often in the form of censorship.

1. In the middle of November 2023, noteworthy examples include the plight of Pustakapias, an independent publisher located in Malang City, which encountered censorship and a subsequent hacking incident on its Instagram account. This occurrence followed the dissemination of a book publication schedule elucidating extraordinary crimes perpetrated by the state during the years 1965-1968. The censorship enacted by Meta purportedly cited fraudulent activities as justification.
2. On 9 December 2023 Butet Kertaradjasa, an artist, received a digital attack in the form of WhatsApp hijacking after holding an art performance with political content in Yogyakarta. Previously, Butet also received a warning from the police not to bring political elements into any performances that were being performed.¹³
3. On 26 November 2023 a digital attack was launched against the website of Charta Politika, a research and survey institution, resulting in a domain incapacitation rendering the site inaccessible. This offensive unfolded subsequent to Charta Politika's dissemination of research findings pertinent to the "Electoral Map Post Constitutional Court Decision & Registration of Presidential Candidates-Vice Presidential Candidates". The published survey delineated that 48% of respondents expressed reservations regarding the suitability of Gibran Rakabuming

¹³ (2023, December 9). Butet Beber Kronologi Nomor WhatsApp Miliknya Dilumpuhkan. *CNN Indonesia*. <https://www.cnnindonesia.com/hiburan/20231209140149-234-1035075/budet-beber-kronologi-nomor-whatsapp-miliknya-dilumpuhkan>

as a candidate for the position of Vice President in the impending 2024 election.¹⁴

7. Attacks Against Journalists

AJI or the Alliance of Independent Journalists was born as a form of the press's retaliation against repressive control from the government during the New Order regime in 1994. The alliance is determined to actively safeguard the freedom of press. Today, AJI continues to provide space and mechanism for journalists to report discrimination they face in the field. Data of case reports received by AJI shows an increase in harassment faced by journalists from 2021 to 2023 (Fig. 18). In 2023, journalists and media experienced a variety of attacks; terror, intimidation, threats, digital attack, destruction/seizure of equipment, criminalization and civil lawsuit, forced deletion of coverage, reporting ban, sexual and physical violence. Spillover attacks on informants are also found in 4 cases. 9 cases of discrimination were reported due to the coverage of political and electoral issues¹⁵.

¹⁴ Dyantoro, S. (2023, November 26). Website Charta Politika Indonesia Kena Gangguan Domain, Tak Dapat Diakses. *Tempo.co*. <https://tekno.tempo.co/read/1801581/website-charta-politika-indonesia-kena-gangguan-domain-tak-dapat-diakses>

¹⁵ N.d. (2024). Bab 4: Makin Berisiko di Isu Lingkungan dan Korupsi, Impunitas Langgengkan Represi. *Aliansi Jurnalis Independen*.

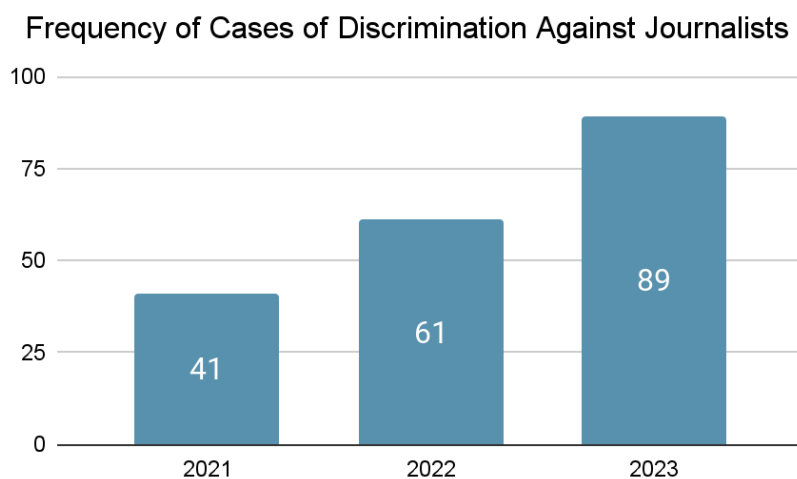


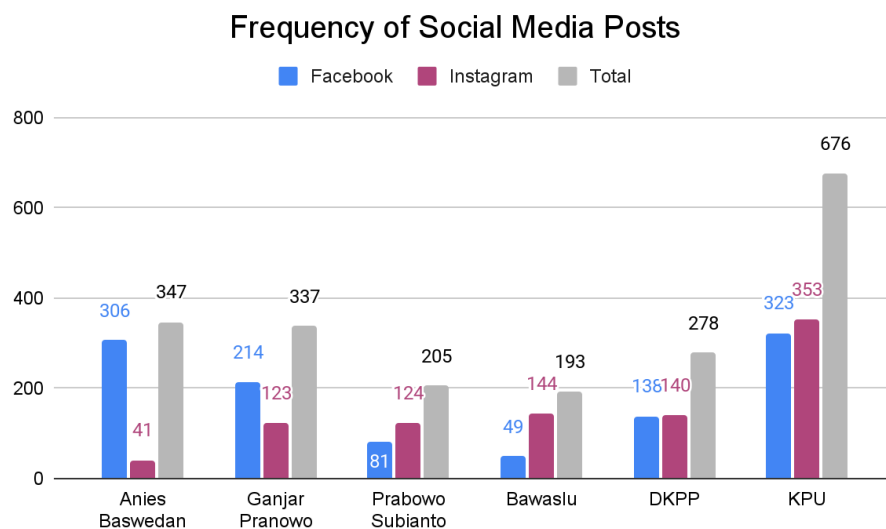
Figure 18. Cases of attacks towards journalists from 2021-2023

Example:

1. [Incident 1](#): On January 5th, 2024, twelve journalists from various media were denied approval to document the sorting and folding process of the voting ballots by the head of Kabupaten Kediri’s branch of KPU (General Election Commission). The head of the regional KPU claimed that access to the process of preparing the ballots is limited to those involved. However, the KPU officials consulted with their colleagues and finally granted the journalists permission to cover the event in a limited manner after a considerable amount of time waiting.
2. [Incident 2](#): On January 16th, 2024, a journalist en route to cover presidential candidate Anies Baswedan’s campaigning activity in Sorong was attacked by a group of 8 people who threw two big rocks at the journalist from behind.

8. The use and choice of social media platforms by presidential candidates

In the period of 23 October to 31 December 2023, the candidates show preference in differing social media platforms (Fig. 19). Presidential candidates Anies and Ganjar were more active on Facebook whereas Prabowo posted more on Instagram. The choice of platform here may indicate the segment of the audience targeted by the different presidential candidates. Data by IDN Research Institute¹⁶ found that Instagram is the main social media platform used by Millennials and Gen Z in Indonesia. On the other hand, Facebook and its community-centric approach is still one of the platforms of choice, despite its declining popularity among younger groups.



¹⁶ N.d. (2024). Indonesian Millennial and Gen-Z Report 2024. *IDN Times*. <https://www.imgs.idntimes.com/>

Figure 19. Frequency of social media posts by candidates and election management bodies

Amongst the Election Management Bodies (EMBs), KPU posted the most information. However, in the period of November 20 to December 10, 2023, only 9.9% of the total posts of the EMBs (39) presented information regarding election procedures or information about elections. With the elections on 14 February 2024, the EMBs should have enhanced the utilization of their social media platforms as the main source of information regarding the voting procedure and the elections.

9. Spending on Advertising on Meta

Since 2020, Meta Ad Library¹⁷ has collected ongoing data regarding advertisements on social issues, elections, or politics on Meta platforms and produced detailed reports regularly¹⁸. Prior to the start of the campaigning period on November 28, the total amount spent on political advertisements on the Meta platforms (Facebook and Instagram) during the period of October

¹⁷ N.d. (2024). Meta Ad Library Report. *Meta*.
<https://www.facebook.com/ads/library/report/?source=onboarding>

¹⁸ Information regarding ads is updated in the Ad Library within 24 hours of the time it gets its first impression. Ads are categorized to be about social issues, elections or politics identified by the content and largely by information regarding who paid for the ad. This may pose a challenge where not all possible ads that are intrinsically political are identified by the system.

5 to November 4 2023, was Rp. 3,780,737,620 across 23,732 ads. Most recently, in the period of November 2, 2023, to January 30, 2024, the total rose to Rp. 23,037,590,258 across 121,907 ads. Political ads in favor of presidential candidates were mostly financed by supporter groups such as volunteering communities, official political campaign accounts, and biased online media.

10. Content and Account Takedown Requests

A number of social media platforms publish reports on takedown requests regularly, every six months. Google and YouTube reports¹⁹ that 225 out of 1,908 items covered by 226 government requests were removed (223 removed due to legal violations, 2 removed due to policy violations). TikTok reports²⁰ that a total of 531 content and accounts were actioned out of 1,198 content and accounts from 224 government requests (229 accounts and content were actioned due to community guidelines violations while 302 were actioned due to local law violations). This data (Fig. 20) suggests that 11.8% of total item takedown requests are actioned by Google and YouTube, whereas 44.3% are actioned by TikTok.

¹⁹ N.d. (2023). Government requests to remove content. *Google Transparency Report*. <https://transparencyreport.google.com/government-removals/government-requests/ID?hl=en>

²⁰ N.d. (2023). Government Removal Requests Report. *TikTok*. <https://www.tiktok.com/transparency/en-us/government-removal-requests-2023-1/>

There appears to be a substantial disparity between the volume of requests by the government and the corresponding actions taken by the platform. Generally speaking, it appears that the majority of government requests are not complied with by either YouTube/Google or TikTok, although TikTok seems to comply more often than YouTube/Google. Furthermore, Meta²¹ does not disclose the number of takedown requests received. However, they do disclose the number of restrictions taken based on local law (3866 content and accounts from Facebook and 1378 from Instagram).

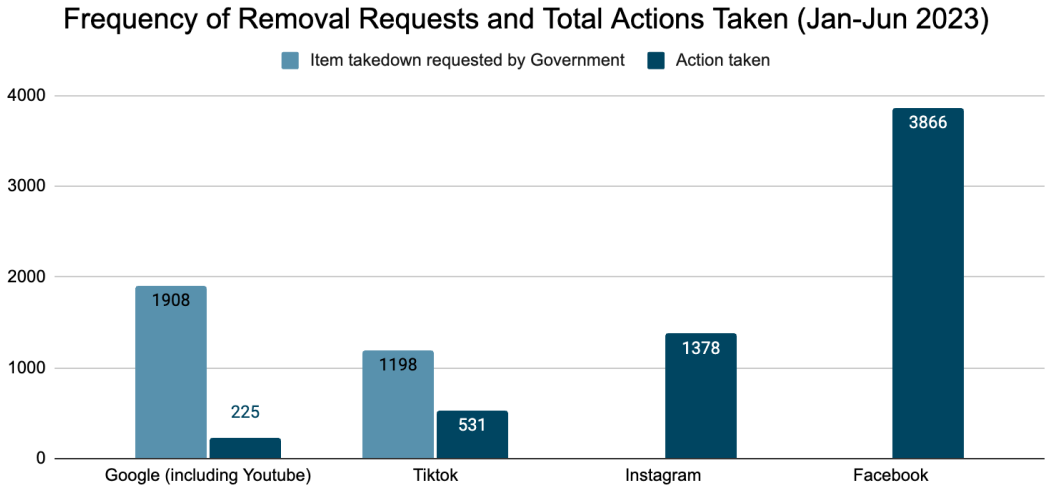


Figure 20. Frequency of item takedown requests and total actions taken in Indonesia in the first half of 2023 based on different social media platforms

²¹ N.d. (2023). Content Restrictions Based on Local Law: Indonesia. *Meta*. <https://transparency.fb.com/reports/content-restrictions/country/ID/>

It is notable that shortly before election day, three prominent accounts commenting on elections reported having received emails from X/Twitter, stating that their tweets/post have received a take-down request by the Ministry of Communication and Information Technology of Indonesia (Fig. 21). The emails stated that the posts violated Indonesian government regulations, although it was not explained what regulations they violated. The emails were sent regarding the following three tweets: [Soe Tjen](#)'s tweet raised doubts regarding vice presidential candidate Gibran Rakabuming Raka, who was observed receiving signals from his campaign team during the presidential debate. Another person, [Andi Kamal Reza](#), received a similar email after expressing skepticism on Twitter/X about the Gerindra party admin on TikTok regarding calls for violence against voters of another presidential candidate, Anies Baswedan. Meanwhile, another account with the user name [@khalidio](#), received a similar email after raising concerns about Joko Widodo, who was suspected of utilizing state resources for campaigning purposes.

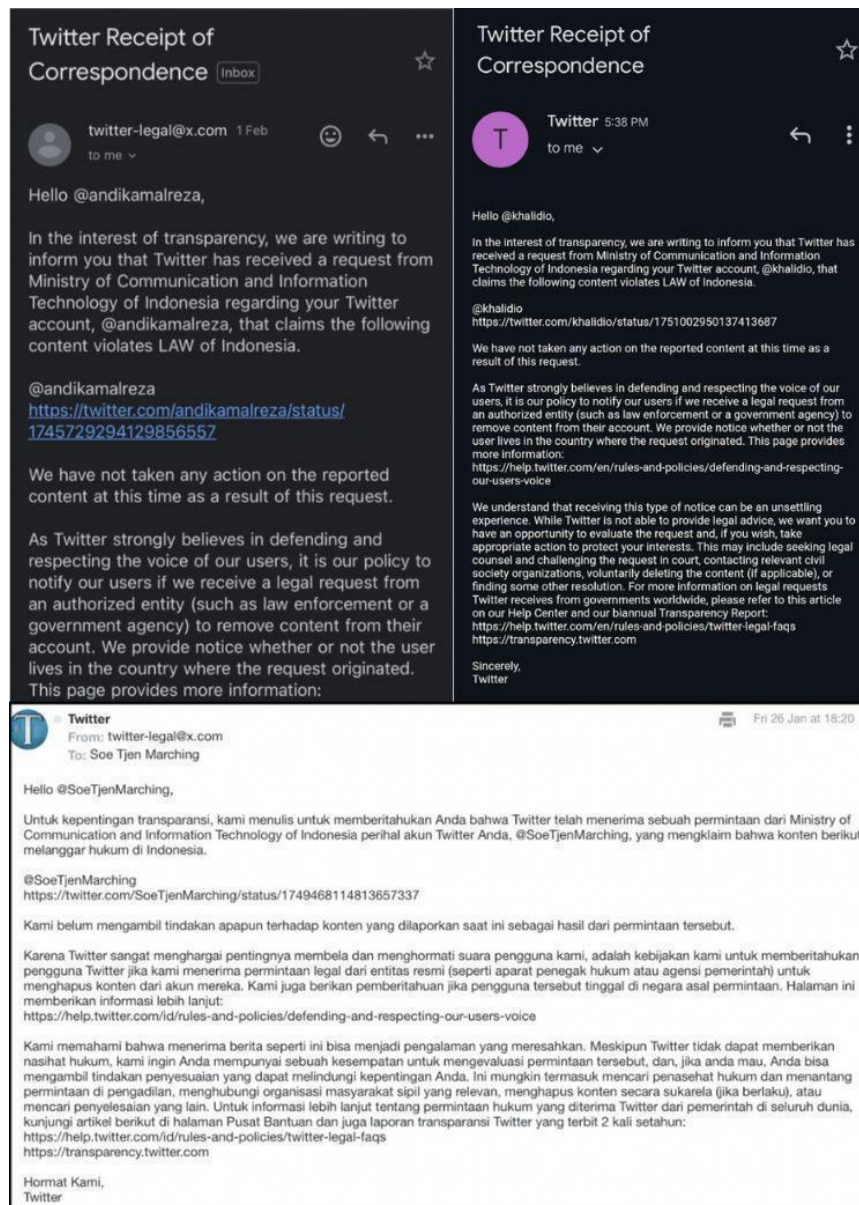


Figure 21. Emails of post-take-down requests by the platform regarding violations of government regulations (Source: Twitter/X)