



# SOCIAL MEDIA MONITORING OF MALAYSIA'S 15TH GENERAL ELECTIONS



Project in partnership with the University of Nottingham Malaysia,  
Universiti Sains Malaysia, and Universiti Malaysia Sabah



University of  
Nottingham  
UK | CHINA | MALAYSIA



USM  
UNIVERSITI SAINS MALAYSIA



UMS  
UNIVERSITI MALAYSIA SABAH

March 2023



Published by:  
Centre for Independent Journalism Malaysia  
L13A-1, Menara Sentral Vista  
150, Jalan Sultan Abdul Samad, Brickfields  
50470 Kuala Lumpur  
Malaysia

Email: [cijmalaysia@gmail.com](mailto:cijmalaysia@gmail.com)  
Website: <https://cijmalaysia.net>  
Facebook: <https://www.facebook.com/CIJ.MY>  
Twitter: [https://twitter.com/CIJ\\_Malaysia](https://twitter.com/CIJ_Malaysia)  
Instagram: [https://www.instagram.com/cij\\_malaysia](https://www.instagram.com/cij_malaysia)  
YouTube: <https://www.youtube.com/cijmalaysia>

The Centre for Independent Journalism (CIJ) is a feminist, freedom of expression watchdog and non-profit organisation that aspires for a society that is democratic, just and free, where all peoples will enjoy free media and the freedom to express, seek and impart information.

Published in Kuala Lumpur in March 2023



# **Social Media Monitoring of Malaysia's 15<sup>th</sup> General Elections**

## **Writers and contributors**

Masjaliza Hamzah

Wathshlah Naidu

Lee Shook Fong

Dineshwara Naidu

Daraneer Balachandar

## **Experts/editors**

Gayathry Venkiteswaran

Professor Zaharom Nain

Dr Ngo Sheau Shi

Dr Mahyuddin Ahmad

Wang Lay Kim

Dr Lee Kuok Tiung

Dr Latif Lai Che Ching

## **Editors**

Jacqueline Ann Surin

Izmil Amri Ismail

## **Layout Designer**

Joelle Majanil

All Rights Reserved © 2023 Centre for Independent Journalism. This report may not be copied or duplicated in whole or part by any means without express prior agreement in writing from CIJ. Some photographs in this report may be copyrighted or the property of others. In such cases, acknowledgement of those copyrights or sources have been given



## ACKNOWLEDGEMENT

We wish to acknowledge the significant contribution of the following persons and organisations towards the monitoring and analysis:

### Experts

Universiti Sains Malaysia: Dr Ngo Sheau Shi, Dr Mahyuddin Ahmad, Wang Lay Kim

University of Nottingham Malaysia: Gayathry Venkiteswaran, Professor Zaharom Nain

Universiti Malaysia Sabah: Dr Lee Kuok Tiung, Dr Latif Lai Che Ching

### Monitors

*Pilot Monitoring:*

24 Monitors from Universiti Sains Malaysia and University of Nottingham, Malaysia

*GE15 Monitoring:*

54 Monitors from Universiti Sains Malaysia, University of Nottingham, Malaysia and the public

### Rapid Response Partners

Architects of Diversity, Asylum Access Malaysia, Beyond Borders, Justice for Sisters, KRYSS Network, North South Initiative, Pusat Komus, Sahabat Wanita, Sisters in Islam and Tenaganita

### Support

National Democratic Institute

### Tech Support

Zanroo Malaysia Sdn. Bhd. (1092951-H)

## DISCLAIMER

The findings and conclusions presented in this report are based on limited data and information available at the time of the monitoring, conducted during the campaign period of the 15th Malaysian General Elections. This report should not be construed as representing a larger population or making generalised claims. Monitoring was conducted with a limited sample size and a limited number of human monitors with limited reach and/or capacity. As such the results may not be applicable to broader populations or contexts. Therefore, caution should be exercised in interpreting and applying the findings of this report beyond the specific parameters of the monitoring guidelines and areas of concern. The authors of this report do not assume any responsibility for any errors, omissions, or misinterpretations that may arise due to the limitations.

# TABLE OF CONTENTS

Executive Summary \_\_\_\_\_ vii

**1. Introduction \_\_\_\_\_ 1**

**2. Monitoring Framework \_\_\_\_\_ 6**

    2.1 Defining hate speech for the monitoring project \_\_\_\_\_ 7

    2.2 Rules for identifying issues, actors and platform \_\_\_\_\_ 8

**3. Methodology \_\_\_\_\_ 11**

    3.1 Methods \_\_\_\_\_ 12

    3.2 Data Collection \_\_\_\_\_ 12

**4. Monitoring Analysis \_\_\_\_\_ 14**

    4.1 Overall highlights \_\_\_\_\_ 14

    4.2 Analysis by issues \_\_\_\_\_ 21

        • Race \_\_\_\_\_ 21

        • Religion \_\_\_\_\_ 25

        • Royalty \_\_\_\_\_ 28

        • Gender and LGBTIQ \_\_\_\_\_ 30

        • Refugees and migrants \_\_\_\_\_ 34

    4.3 Coordinated Inauthentic Behaviour (CIB) \_\_\_\_\_ 38

    4.4 Analysis of social media platforms \_\_\_\_\_ 41

**5. Recommendations \_\_\_\_\_ 45**

**Annex I \_\_\_\_\_ 49**



## Executive Summary

After the historic 2018 elections, Malaysia was suddenly thrown into political uncertainty in early 2020 following the defections of several government members of parliament (MPs). This resulted in three prime ministers leading three governments after the 14<sup>th</sup> general election. With so much political instability, the stakes for the 15<sup>th</sup> General Elections (GE15) in 2022 were high and social media inevitably became an important space for power brokers and politicians.

Social media is a popular space for Malaysians to interact with each other. Since the 2008 elections, politicians, political parties and their supporters have increasingly turned to social media to engage with netizens. It has since become an important channel to shape political discourse, with politicians and opinion leaders bypassing the mainstream and online media to post comments and statements directly on various platforms.

In the run-up to previous elections, online users employed divisive language and hate-based narratives around race, religion and royalty, popularly known as 3R. The GE15 was expected to be no different. This was especially because a couple of issues had made social media an important space that shapes political discourse, including by promoting hate and toxic rhetoric. One was the political uncertainty of the past five years. This meant the stakes were high in GE15. The other was the rollout of automatic registration for voters above 18. This meant there were 1.4 million voters<sup>1</sup> aged between 18 to 20, whose voting patterns were unknown but whose use of social media was well-documented.

Hence, the Centre for Independent Journalism (CIJ), in collaboration with the University of Nottingham Malaysia (UNM), Universiti Sains Malaysia (USM) and Universiti Malaysia Sabah decided to conduct a social media monitoring project pre-, during and post-GE15.

A key objective of this monitoring was to identify the severity levels of hate speech and expose the sources and patterns of hateful and toxic posts on social media, specifically Facebook, Twitter, YouTube and TikTok. While the politicians themselves may not be the originator of hateful posts or false information, they did not curb or censure followers who responded with hate speech and disinformation. For some politicians, their repeated use of pejorative terms against their opponents pointed to them being aware of this pattern of behaviour and the kind of response it generated.

The monitoring looked at levels of hate speech<sup>2</sup> that were distributed by actors such as politicians, political parties, government agencies, media organisations, and key opinion leaders (KOLs) around several key topics. These were the 3R, gender and lesbian, gay, bisexual, transgender, intersex and queers (LGBTIQ) persons, and refugees and migrants. The monitoring also identified potential Coordinated Inauthentic Behaviour (CIB) such as bots and cybertroopers.

CIJ developed a severity scale in which it categorised the levels of hate speech as follows: Level 1 disagreement / non-offensive; Level 2 offensive / discriminatory; Level 3 dehumanising / hostile; and Level 4 Incitement / call for action (violence). The team adapted the thresholds in the Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility

<sup>1</sup> <https://www.nst.com.my/news/politics/2022/11/848134/ge15-young-voters-cast-kingmakers>

<sup>2</sup> Article 19's definition of hate speech: Any expression of discrimination and advocacy of hatred towards individuals or groups on the basis of specific protected characteristics. These expressions, opinions or ideas of hate may be written, non-verbal, visual, artistic or in some other form, and may be disseminated through any media, including the internet, print, radio, or television.

or violence<sup>3</sup> and used the 3M framework<sup>4</sup> of Messenger, Message and Messaging to analyse the levels of hate speech.

Social media was mined using a customised automated tool which scraped data on identified accounts, using a combination of 544 keywords and character embeddings. CIJ-trained human “monitors” then reviewed, categorised and tagged the data.

CIJ carried out a pilot study from 16 August to 30 September 2022 to test the tool and review the data in preparation for use during GE15. After Parliament was dissolved on 10 October 2022, the final monitoring was carried out between 20 October to 26 November.

Key findings include:

1. PAS president Hadi Awang and his party were the biggest amplifiers of race. For example, Hadi’s TikTok feed claiming the DAP was merely using Malay candidates to gain voter traction had 2.5 million engagement<sup>5</sup> counts on TikTok, the highest across all social media platforms. Posts on race were also found to perpetuate disinformation. Further, after election day, content created by young TikTok users to manufacture fear went viral and had high cross-platform amplification. Content on race frequently intersected the religion.

2. The monitoring saw a shift in focus where religious narratives became the main divisive and polarising tool, as it clearly intersected with race and became the basis for the attacks on LGBTIQ persons. There was high cross-platform amplification, including through posts by key opinion leaders (KOL) and other influencers. One was by famous singer Jamal Abdillah who called on voters to reject the DAP which if voted into power, would purportedly restrict the azan (call to prayer) and tahfiz (religious) schools. And the other was by actor Zul Huzaimy Marzuki whose wish to “slaughter infidels” was only prevented because of the law.

3. The targeting of refugees and migrants on social media did not seem to be part of any coordinated political campaign. However, it ranked highest for severity at Level 4. Posts about these two communities contained explicit suggestions calling for physical harm, damage or death. The dehumanisation of these two marginalised groups was made worse when the Immigration Department asked social media users to complain and submit information about people they suspected were without documents. It is however noted that the attacks against refugees and migrants on social media have been consistent and in fact continued post elections<sup>6</sup>.

4. Hate speech against the LGBTIQ community was, by contrast, clearly a political tool, often coupled with terms such as “liberal” and “anti-Islam”, and used largely to discredit Pakatan Harapan (PH), the Democratic Action Party (DAP) and Anwar Ibrahim. At the same time, body shaming and the policing of modesty were aimed at women from across the political spectrum.

5. Posts on royalty, which were low during much of the monitoring period, spiked after polling day when it became clear the institution would be a key player in deciding who the next government would be. Most posts under royalty were not attacks against the royal

<sup>3</sup> Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence, 11 January 2013, A/HRC/22/17/Add.4

<sup>4</sup> Adopted from the “Monitoring of Media Coverage of Elections: Toolkit for civil society organisations”, Council of Europe, November 2020. 3M refers to Messenger, Message and Messaging. The Messenger is the actor (source of information, for example, political party and or leader, media outlet, or influencer) who spreads the Message (topics and narratives disseminated by actors). Messaging refers to how a message travels.

<sup>5</sup> Engagement count includes, ‘views’, ‘likes’, ‘shares’ and ‘comments’ on the platform.

<sup>6</sup> Geng IC Merah jadi duri dalam daging bawa masuk PATI

<https://www.kosmo.com.my/2023/03/15/pemastautin-tetap-jadi-duri-seludup-pati/>



institution but against other actors. One example was the attack on Muhyiddin Yassin for “disloyalty” to the king in rejecting the royal advice to form a unity government.

6. The automated system and monitors identified CIBs during the monitoring period. However, these did not have any significant impact on shaping narratives about the issues that were monitored. User-generated comments (UGCs) from regular social media users were the biggest amplifiers.

The monitoring has sparked interest and debate on the key role of the State in combating hate speech on social media. The virality of simplistic and inaccurate TikTok content, for example the 13 May videos post-elections or Hadi’s video accusing the DAP of using Malay candidates has also brought to the forefront the question of social media accountability, especially in relation to the standards and processes they adopt in moderating content. There is a critical need to interrogate the use of social media as a key tool in shaping the political and electoral discourse and what further standards and measures are needed moving forward within this technology driven environment.



# SOCIAL MEDIA MONITORING OF MALAYSIA'S 15TH GENERAL ELECTIONS



## 1 | INTRODUCTION



# 1. INTRODUCTION

Malaysia experienced political uncertainty for three years after the reformist Pakatan Harapan (PH) government fell in early 2020 following defections by Members of Parliament (MPs) from its ruling coalition. By March 2020, Muhyiddin Yasin from Parti Pribumi Bersatu Malaysia (Bersatu) was appointed the new prime minister. He led a cobbled-together coalition, Perikatan Nasional (PN). This coalition included United Malays National Organisation (Umno), which is riddled with corruption scandals, and the right-wing Islamist party, Malaysian Islamic Party (PAS). Eventually, Muhyiddin's government fell and Ismail Sabri was appointed prime minister in August 2021. This brought Umno back into power. Throughout this political upheaval, Malaysia was also navigating COVID-19.

During this time, freedom of expression and information integrity in Malaysia was on the backslide, with increasing weaponization of repressive laws to harass, intimidate and silence dissenting voices. The government declared a national emergency and suspended Parliament in January 2021. It then enacted a draconian law which carried a fine of up to RM100,000 and up to three years imprisonment, ostensibly to combat fake news.

On the global stage, Malaysia succeeded in being elected as a United Nations Human Rights Council member. But at home, the government continued to use overly-wide and arbitrary laws such as the Sedition Act and the Communications and Multimedia Act (CMA) to arrest and investigate critics and dissenters, including critical netizens. Investigations against human rights activists and media also intensified during this period. Activists were frequently investigated and even charged for speaking out and acting against human rights abuses.

Meanwhile, Ismail Sabri's administration and the opposition PH coalition signed a memorandum of understanding, promising selected parliamentary and electoral reforms in return for support of the government's budget. At the same time, Umno and its coalition party members experienced overwhelming success at the Melaka and Johor state elections in 2021 and 2022 respectively. This further cemented the return of Umno/Barisan Nasional (BN) to power. Against this backdrop, the 15<sup>th</sup> General Elections (GE15) was expected to be another game changer.

Based on what happened in previous years, the expectation was that GE15 would be preceded by a concerted effort to spread divisive language and hate narratives based on specific political agendas. These narratives would be grounded in 3R and would mostly target particular communities at risk. The situation was expected to be made worse by a lack of government transparency and denial of access to information, which creates an environment conducive to corruption and rights infringement.

In this climate, a strong information ecosystem is crucial for holding the state accountable and for citizens to openly criticise the government and expose wrongdoing. However, the current information and media landscape is highly politicised and often used by the ruling government, as well as numerous political factions, to advance their respective propaganda. Further, the media is under digital siege as social media platforms become the main source of information.

At the same time, the rapid advance in digital infrastructure and awareness has created broader spaces for engagement as well as wider dissemination. On the other hand, there is also the potential for a wider digital divide and an emergence of new threats. These include the proliferation of disinformation, negative narratives, and hate speech; an increase in digital surveillance; privacy violations; data engineering; online gender-based violence;

the emergence of cybertroopers; and manipulation of algorithms that facilitate echo chambers and filter bubbles; and online bullying. These threats accelerated divisions in society and political instability, and were estimated to hugely influence netizens and voters leading up to and during GE15.

### Legal framework against hate speech

One of this project's key objectives was to identify hate speech on social media, including their sources and severity level during the monitoring period. Where appropriate, a response from civil society would be crafted. This was in line with CIJ's mandate to promote freedom of expression according to a normative human rights framework and to make the state accountable in providing a conducive environment for free speech, media freedom and the protection of rights within digital spaces. CIJ's media monitoring exercises have in the past recorded valid forms of critical, dissenting and satirical expressions, as well as serious violations such as false claims, threats, hate speech and other harmful narratives.

### Legal guarantee of freedom of expression in Malaysia

Article 10 of Malaysia's Federal Constitution grants citizens the right to freedom of speech, freedom of assembly and freedom of association, while Article 8(2) prohibits discrimination based on religion, race, descent, place of birth, or gender.

### Criminalising hate speech under Malaysia's current legal framework

However, instead of applying these protective provisions, the state uses oppressive legislations such as the Sedition Act, Section 233 of the CMA, and Sections 504 and 505 of the Penal Code to silence dissenting opinions and expressions. Infringements by State and non-State actors are well documented by local civil society organisations (CSOs) such as CIJ and Suaram, as well as international CSOs.

Section 4A of Malaysia's Election Offences Act (1954) makes it an offence to promote ill will or hostility. Section 9 of the same Act prohibits direct or indirect undue influence through the use of force, violence, or restraint, or by inflicting injury, damage, harm or loss. Both sections are specific to ensuring the free exercise of a citizen's electoral right. However, these sections offer little security for non-citizens and marginalised 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 its use – apart from self-moderation by users – in the hands of the platforms themselves or the Malaysian authorities. This puts a lot of faith in profit-making companies, whose recent downsizing has also affected its content-moderating roles, and in a government which has tended to be heavy handed or selective in its response.

Inflammatory speech especially on race and religion is a real issue in multi-ethnic and multi-religious Malaysia and government proposals to address them centre around more legislation, including proposing the enactment of an Anti-Discrimination Bill, the National Harmony and Reconciliation Commission Bill, and the Racial and Religious Hatred Bill. Instead of enacting more criminal provisions, civil society has called on the Malaysian government to consider "positive policy initiatives to address discrimination and conflict in society and to promote tolerance and intercultural understanding".

### International standards in determining hate speech

There is no universally-accepted definition of hate speech.



Articles 19 and 20 of the International Covenant on Civil and Political Rights (ICCPR) address State obligations in the following contexts:

- *Prohibition of hate speech*

Article 20(2) obliges states to prohibit by law severe forms of hate speech that “advocates national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence”, in accordance with requirements under Article 19(3).

- *Three-part test in prohibiting hate speech*

A state may impose restrictions to freedom of expression only in limited and exceptional circumstances and is required to meet the three-part test in accordance with Article 19(3) of the ICCPR:

- Legality: must be provided for by law
- Legitimacy: must pursue a legitimate aim, such as respect for the rights of others
- Necessity and proportionality in a democratic society

### *Rabat Plan of Action*

The Rabat Plan of Action, adopted in 2012, provides a high threshold for restricting freedom of expression with its six-part test on “incitement to hatred”, in addition to the three-part test under Article 19(3) of the ICCPR. The six-part threshold test considers (1) context of the statement, (2) status of the speaker, (3) intent to incite violence against a target group, (4) content and form of statement, (5) extent of its dissemination, and (6) likelihood of harm, including imminence.<sup>7</sup>

### Social media platforms addressing hate speech

Social media platforms define hate speech in their terms of service or community standards. In May 2016, Facebook, Twitter and YouTube signed a code of conduct with the European Commission on regulating illegal hate speech online.<sup>8</sup>

a. Facebook’s Community Standards define hate speech as a direct attack against people on the basis of what they call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease.<sup>9</sup>

b. Twitter’s Hateful Conduct Policy sanctions users whose tweets or direct messages:

- i. make violent threats against a target
- ii. wish, hope or call for serious harm on a person or group of people
- iii. make references to mass murder, violent events, or specific means of violence where protected groups have been targets or victims
- iv. incite against protected categories

<sup>7</sup> “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](https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/SeminarRabat/Rabat_threshold_test.pdf)

<sup>8</sup> “The Code of conduct on countering illegal hate speech online”, European Commission, 22 June 2020, [https://ec.europa.eu/commission/presscorner/detail/en/qanda\\_20\\_1135](https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1135)

<sup>9</sup> “Facebook Community standards: Hate speech”, Meta, Available at [https://www.facebook.com/communitystandards/hate\\_speech/](https://www.facebook.com/communitystandards/hate_speech/) cited in Papcunová, J., Martončík, M., Fedáková, D. et al. Hate speech operationalization: a preliminary examination of hate speech indicators and their structure. *Complex Intell. Syst.* (2021). <https://doi.org/10.1007/s40747-021-00561-0>

v. repeat non-consensual slurs, epithets, racist and sexist tropes, or other content that degrades someone

vi. make hateful imagery<sup>10</sup>

c. Hate speech on YouTube is identified as content that “incites hatred or violence against groups based on protected attributes such as age, gender, race, caste, religion, sexual orientation, or veteran status”.<sup>11</sup>

d. TikTok does not allow “any hateful behaviour, hate speech, or promotion of hateful ideologies”. This includes content that attacks a person or group because of protected attributes, including: caste, ethnicity, national origin, race, religion, tribe, immigration status, gender, gender identity, sex, sexual orientation, disability and serious disease. The Community Principles also define hate ideologies as “systems of beliefs that exclude, oppress, or otherwise discriminate against individuals based on their protected attributes, such as racial supremacy, misogyny, anti-LGBTQIA+, and antisemitism”<sup>12</sup>.

---

<sup>10</sup> “Hateful Conduct”, Twitter, February 2023, <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy>

<sup>11</sup> “How does YouTube protect the community from hate and harassment?”, Our Commitments: Hate versus harassment, YouTube, <https://www.youtube.com/howyoutubeworks/our-commitments/standing-up-to-hate/>

<sup>12</sup> Community Guidelines, TikTok, Last updated March 2023, <https://www.tiktok.com/community-guidelines/en/?c-gversion=2023>

# SOCIAL MEDIA MONITORING OF MALAYSIA'S 15TH GENERAL ELECTIONS



## 2 | MONITORING FRAMEWORK



## 2. MONITORING FRAMEWORK

Based on observations from past election monitoring projects, CIJ anticipated an increase in hate speech, especially on social media during GE15. These tactics have been frequently used to control narratives and influence public understanding of key issues. Hate speech also distracts and diverts attention away from problem-solving, apart from increasing the potential for harm against individuals and marginalised communities.

This project's overarching goal was to strengthen Malaysian society's resilience against hate-based narratives prior to and during GE15. Its goal was also to defend and promote freedom of expression and information integrity in Malaysia as essential components in a democracy.

The project's monitoring framework was grounded in human rights. That is, recognising peoples' right to political participation and respect for their civil and political rights of freedoms of expression, opinion, and assembly; right to information; and the right to effective legal remedy for electoral violations within the wider intersections of race, gender, religion, sexual orientation, gender identity and nationality.

The monitoring framework was also guided by the principles of impartiality as well as systematic, comprehensive, and accurate collection and analysis of information.

### Scope of monitoring

The project employed the 3M method in its analysis of:

- a. Toxic narratives and hate speech around 3R, gender and LGBTIQ, refugees and migrants
- b. Inauthentic actors (coordinated or lone), disinformation and manufactured amplification of toxic narratives and hate speech around 3R, gender and LGBTIQ, refugees and migrants.

Using an automated tool, we looked at divisive and toxic narratives on select actors' official Facebook, Twitter, YouTube and TikTok accounts. The monitors also checked for CIB.

### Scope of analysis

The monitoring dataset was analysed according to:

- Content analysis: message or narrative of the post
- Network identification and analysis: mapping actors and their relationships

### 2.1 Defining hate speech for the monitoring project

There is no universally-agreed definition of hate speech, including in Malaysia, or of levels of hate speech. As a result, the monitoring project adapted its own levels. In this monitoring, the severity level of hate speech may not necessarily meet the high threshold in the Rabat Plan of Action. Nonetheless, the severity levels address the progression of speech that is intolerant, discriminatory and dehumanising, and that incites violence and bodily harm.

The monitoring considered 2 basic elements of hate speech<sup>13</sup>:

- Hate: “the intense and irrational emotion of opprobrium, enmity and detestation towards an individual or group, targeted because of their having certain – actual or perceived – protected characteristics (recognised under international law<sup>14</sup>)”
- Speech: “any expression imparting opinions and ideas – bringing an internal opinion or idea to an external audience. It can take many forms: written, non-verbal, visual or artistic and can be disseminated through any media including internet, print, radio or television”.

The project established parameters of hate speech and categorised speech on a spectrum of four severity levels with corresponding characteristics:

- Level 1: disagreements or non-offensive language*
- Level 2: offensive or discriminatory language*
- Level 3: dehumanising or hostile language*
- Level 4: causing incitement or calls for violence*

It further adapted the threshold test in the Rabat Plan of Action, and the categorisation using the 3M framework into the following table:

*Table 1: Threshold test for hate speech severity*

MESSENGER	1. Speaker	• Position/status (degree of influence) in society.
	2. Intent	• Intent to incite hatred or inflict harm.
MESSAGE	3. Context	• Locating the speech within social, religious and political context, and the power dynamics at the time the speech was made and disseminated. This may include past historical context, policies, or social norms.
	4. Content & Form	• Degree to which speech was provocative and direct. • Form, style and nature of arguments used in the speech.
MESSAGING	5. Audience	• Reach of speech, its public nature, magnitude and audience size. • Frequency, quantity and extent of communications. • Whether the statement is circulated in a restricted environment or widely accessible to the public. • Whether the audience had the means to act on the incitement.
	6. Medium	• Means of messaging and amplification (one platform or multiple or cross-platforms). • Use of inauthentic accounts (bots, cybertroopers).

## 2.2 Rules for identifying issues, actors and platform

### a. Selection of Actors

The actors monitored were chosen from the following categories:

#### 1. Political parties

All key political parties were included, taking into account the communities they represent, their geographical representation and influence on social media.

<sup>13</sup> 'Hate Speech' Explained: A Toolkit, 2015 Edition, Article 19 <https://www.article19.org/data/files/medialibrary/38231/%27Hate-Speech%27-Explained---A-Toolkit-%282015-Edition%29.pdf>

<sup>14</sup> See Camden Principles on Freedom of Expression and Equality, “Hate Speech’ Explained: A Toolkit, 2015 Edition”, Article 19, 2015.

## 2. Politicians

Politicians were chosen to represent all key political parties, taking into account their engagement on issues covered under this monitoring project, gender and geographical representation, influence on social media, and whether they had a track record of spreading toxic or positive narratives.

## 3. Media

Top media brands were chosen based on ratings from Statista<sup>15</sup> and the Reuters Institute Digital News Report<sup>16</sup>, their social media engagement rankings, as well as language and geographical representation.

## 4. Key government agencies

Agencies were chosen based on their engagement level or if they were the source of information on issues covered under this monitoring project.

## 5. Key Opinion Leaders (KOLs)

KOLs were chosen based on their influence on social media or if the accounts had a track record of spreading disinformation, toxic narratives, or positive comments that gained traction.

The project also looked at inauthentic actors – whether these were human cybertroopers or bots, coordinated or lone – and identified their impact in influencing discourse on the five issues.

### **b. Selection of issues**

The project monitored the levels of hate speech surrounding narratives and content around 3R, gender, refugees and migrants, and LGBTIQ persons that were distributed by the above-named actors.

CIJ and civil society's experience in previous monitoring exercises found that these issues have increased salience during elections.

Keywords and sub-themes for each issue were identified in the three main languages of Malay, English and Mandarin.

### **c. Selection of platform**

Using Social Media Monitoring (SMM) tools, we looked at divisive and toxic narratives on official Facebook, Twitter, YouTube and TikTok accounts.

The platforms were selected based on the top social media ranking in Malaysia<sup>17</sup>, where in 2021, 47% of Malaysians share news via social media, messaging or email<sup>18</sup> as follows: Facebook (56%), YouTube (36%), Instagram (22%) and Twitter (16%). Tiktok was identified

<sup>15</sup> Share of respondents who consume news through online media in Malaysia in 2022, by media brand", Statista, June 2022, <https://www.statista.com/statistics/982709/malaysia-weekly-use-online-media-by-brands/>

<sup>16</sup> Zaharom Nain, "Malaysia", in Reuters Digital News Report 2021: 10<sup>th</sup> Edition, Reuters Institute for the Study of Journalism, pp 140-141

<sup>17</sup> Zaharom Nain, "Malaysia", in Reuters Digital News Report 2021: 10<sup>th</sup> Edition, Reuters Institute for the Study of Journalism, pp 139-140, [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital\\_News\\_Report\\_2021\\_FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf)

<sup>18</sup> Reuters Digital News Report 2021: 10<sup>th</sup> Edition, Reuters Institute for the Study of Journalism, [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital\\_News\\_Report\\_2021\\_FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf)



as priority as it ranked the second top mobile app in Malaysia as of 2022, and the majority of its users are under the age of 30<sup>19</sup>. TikTok had also significantly impacted the outcome of the Philippines election in May 2022<sup>20</sup>.

#### ***d. Rules in determining inauthentic accounts***

The following rules were used in determining if a post was created by an inauthentic account<sup>21</sup>:

1. If an account was recently activated, in particular around the elections
2. If an account had limited followers
3. If an account had similar type of messages and responses
4. If the account was following similar type of accounts and reposting similar content
5. The speed of dissemination and response

Refer to Annex I for accounts monitored under this project.

---

19 "Social Media Penetration in Malaysia [Research]", Digital Business Lab, 26 July 2022, <https://digital-business-lab.com/2022/07/%E2%91%A1-social-media-penetration-in-malaysia-research/>

20 De Guzman, C. (2022, May) A Dictator's Son Rewrites History on TikTok in His Bid to Become the Philippines' Next President. Time Magazine. Link: <https://time.com/6173757/bongbong-marcos-tiktok-philippines-election/>. Accessed on 6 April 2023.

21 Twitter accounts were also cross checked against Botometer. Botometer investigates a Twitter account's activity and gives it a score. Higher scores indicate more bot-like activity. Link: <https://botometer.osome.iu.edu/>

# SOCIAL MEDIA MONITORING OF MALAYSIA'S 15TH GENERAL ELECTIONS



## 3 | METHODOLOGY

## 3. METHODOLOGY

### 3.1 Methods

Two monitoring exercises were conducted. The pilot was conducted prior to the dissolution of Parliament (15 August – 30 September 2022), and the actual monitoring began after the Election Commission's announcement on 20 October till 26 November 2022.

The pilot was designed to test the monitoring tool and review the validity of the data collected. The pilot also served to assess the readiness and reliability of the monitors, as well as the efficacy of the monitors' reporting structure.

The actual GE15 monitoring was conducted before, during and after GE15 and included monitoring on nomination day on 5 November and on election day on 19 November. When the tight contest ended in a hung parliament, CIJ made the decision for data collection to continue for an additional week.

#### 3.1.1 Literature review

We conducted a literature review of hate speech and disinformation during elections, in particular using local and international election monitoring case studies.

We also met with allies who had practical experience of media monitoring to learn about their monitoring framework, scope, tools, insights, and practical tips and recommendations for developing and implementing this monitoring framework.

#### 3.1.2 Data collection methods

##### a) Using tools for data collection

CIJ contracted a Malaysian data technology company Zanroo which used an Application Programming Interface and search algorithms from within Facebook, Twitter, YouTube and TikTok to collect quantitative data by keywords and specific pages. This enabled the scanning and processing of millions of content on these platforms for the "scraping" of data that fit the keywords and accounts specified by CIJ and its team of experts.

##### b) Engaging human monitors for processing the collected data

CIJ put together a team of monitors – mostly university students – who were then trained to review, categorise and tag the dataset according to key words and the severity level of the speech. They also reported daily CIJ to identify potential rapid responses that were needed from civil society.

### 3.2 Data collection

The different types of posts collected under the monitoring were categorised under these definitions:

- **Total processed posts:** This is the dataset used for the analysis. It is broken down under these headings: actor, issue and severity of speech. Each heading has its own figure.
- **Own post:** Refers to an original post by an actor. Some actors have own posts that did not fit the criteria but generated comments and mentions that fit the criteria.



- **UGC**s: User-generated comments refer to comments/messages by non-actors posted under the original posts, as well as mentions under other actors.
- **Mention**: Refers to posts where actors are mentioned or tagged related to an issue, be it within their own posts or in UGCs.
- **Message**: An actor may mention several issues in one post that may appear under different templates. If one single post is replicated in three templates (for example, under race, religion, etc), it will count as three data points.

### 3.2.1 Intercoder reliability

During the final phase, an intercoder reliability test was conducted to ascertain the reliability of human monitor tagging for severity in the dataset. The test used samples of severity levels to determine intercoder reliability. The Fleiss' kappa statistical model was chosen because it allowed for measurement of reliability of more than three coders. A team of 15 monitors took part in the reliability test and their score was 0.459, which fell within the moderate range of between 0.41 to 0.60. Given the magnitude of data as well as the complexity of the issues, this value was considered reliable enough.

# SOCIAL MEDIA MONITORING OF MALAYSIA'S 15TH GENERAL ELECTIONS



## 4 | MONITORING ANALYSIS

## 4. MONITORING ANALYSIS

### 4.1 Overall highlights

The final dataset for our analysis was derived through a three-part filtering process:

	ACTION	METHOD	UNIQUE MESSAGES
Step 1	Initial filtering of all identified actors, including own posts and UGCs.	Automated Tool	2,323,427
Step 2	Second level of filtering by actors and the identified issues, including own posts and UGCs.	Automated Tool	274,481
Step 3	Final filtering to discard irrelevant posts that were not relevant to the severity criteria.	Human Monitors	99,563

The final number analysed was 99,563 unique messages<sup>22</sup>, which was 4.3% of the initial posts scraped by the automated tool. Data showed that the bulk of posts originated from regular UGCs (62,079) and posts that mentioned specific actors (33,381), rather than the identified actors' own posts (2,433).<sup>23</sup>

Race-based narratives topped the chart with 66,933 posts. This was followed by religion with 24,484; royalty 14,320; gender and comments targeting the LGBTIQ people 7,596; and comments targeting migrants and refugees being the lowest at 3,819.

*Table 2: Total Posts by Issues*

ISSUES	TOTAL POSTS
Race	66,933
Religion	24,484
Royalty	14,320
Gender and LGBTIQ	7,596
Refugees and migrants	3,819
<b>Total Posts</b>	<b>117,152<sup>24</sup></b>

Overall, Levels 3 and 4 posts only made up 0.15% of the total unique messages monitored. It is worth noting that none of the monitored actors' own posts were at Level 4. It was largely UGCs which amplified the hate.

<sup>22</sup> Unique messages are distinct values of the total number of posts reviewed.

<sup>23</sup> Data is not mutually exclusive to a specific type of post, thus the total type of posts may vary from the distinct value of total posts (99,563).

<sup>24</sup> Data is not mutually exclusive to a specific issue; thus, the total of issues (99,563) may vary from the distinct value of total posts (99,563).



**Table 3: Total Unique Messages by Severity Level**

TAG	UNIQUE MESSAGES
Level 1 disagreement / non-offensive	81,763
Level 2 offensive / discriminatory	17,656
Level 3 dehumanising / hostile	105
Level 4 incitement / call for violence	39
<b>Total unique messages</b>	<b>99,563</b>

The majority of 81,763 posts during the monitoring was at Level 1, (for own posts + UGCs + mentions) and categorised as differences of opinion and non-offensive in nature. Meanwhile, a total of 17,800 of all posts (18% or almost one-fifth) were classified as between Levels 2 and 4. This is worrying as it shows that on social media, instead of accepting differences of opinion, one in five users responded with offensive, discriminatory, dehumanising or hostile comments, some of which incited violence or called for bodily harm.

**Table 4: Severity by Issues (Total Processed Posts)**

ISSUES	Level 1 disagreement non-offensive	Level 2 offensive discriminator	Level 3 dehumanising hostile	Level 4 incitement call for violence	TOTAL POSTS
Race	54,827	12,054	47	5	66,933
Religion	21,158	3,315	8	3	24,484
Royalty	13,375	922	23	0	14,320
Gender and LGBTIQ	4,007	3,580	9	0	7,596
Refugees and migrants	2,930	831	26	32	3,819
<b>Total Posts</b>	<b>96,297</b>	<b>20,702</b>	<b>113</b>	<b>40</b>	<b>117,152</b>

#### 4.1.1 Actors fuelling narratives

The number of own posts was low (3,765), with the bulk originating from UGCs (105,698), followed by mentions (68,628).

**Table 5: Issues by Type of Posts**

	METHOD	OWN POSTS	MENTION	TOTAL
Media	7,811	1,371	20,212	29,394
Government	1,550	152	2,892	4,594
Key Opinion Leaders (KOL)	1,178	202	4,774	6,154
Political Parties	21,808	879	30,033	52,720
Politicians	36,281	1,161	47,787	85,229
<b>Total</b>	<b>68,628</b>	<b>3,765</b>	<b>105,698</b>	<b>178,091</b>

Politicians were the key actors contributing to race narratives. This was followed by messages from media actors on religion, royalty, gender and targeting LGBTIQ persons. The government was the main actor who targeted refugees and migrants.

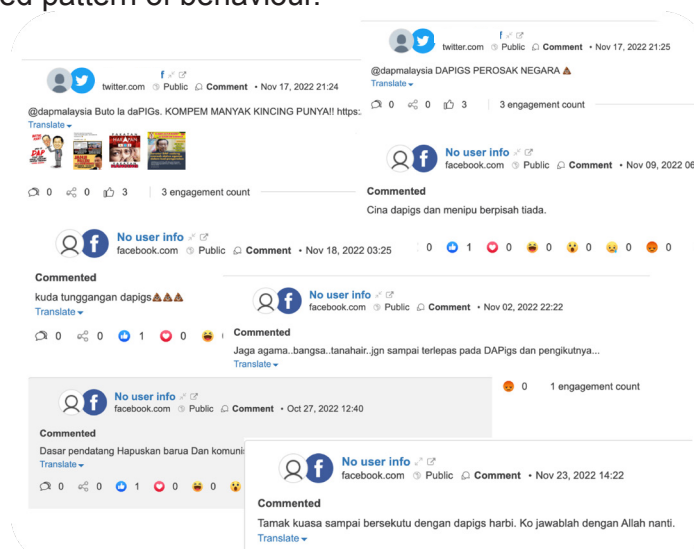
Compared with other actor categories, the media generated the highest number of own posts (1,371), given that they were posting their own news content on the different platforms.

None of the own posts from the actors monitored reached Level 4. Hadi was the only actor with two own posts that reached Level 3. A number of other actors, including Muhyiddin, Lim Guan Eng, Syed Saddiq, Dr Mahathir Mohamad, Zahid Hamid, the Immigration Department of Malaysia, DAP, PAS, Bersatu and (Malaysian United Democratic Alliance (MUDA) had one Level 3 post each.



Facebook, 22 October 2022

We also observed that posts from some key actors became the launching pad from which hate messages were amplified by social media users. There was consistent use of a few terms – for example, “pengkhianat” (traitor) and “DAPigs” – that functioned as triggers for the most severe attacks. The monitoring showed that most of the hate messages did not originate from the actors’ own posts but from UGCs and mentions. Posts about different issues and on different social media platforms also used these terms repeatedly, pointing towards a concerted pattern of behaviour.



Facebook and Twitter, 2-23 November 2022

The main actors monitored did not censure or curb commenters from spewing hatred. They did not stop their supporters from spreading extreme messages. Neither did they utilise the safeguard and filter features offered by social media platforms to restrict certain keywords. In other words, the actors enabled hate speech.

Table 6: Severity by Type of Posts

SEVERITY LEVEL	Mentions	Own Posts	UGCs	TOTAL POSTS
Level 1	29,979	2,429	61,682	94,090
Level 2	8,486	280	12,760	21,526
Level 3	25	4	81	110
Level 4	2	0	32	34
<b>Total Posts</b>	<b>38,492</b>	<b>2,713</b>	<b>74,555</b>	<b>115,760</b>

4.1.2 Perpetuation of disinformation

Some actors with high numbers of own posts and/or UGCs perpetuated disinformation without giving evidence or by bending the truth and/or falsifying information.

Actors such as PAS and Hadi, for example, repeatedly accused the DAP of being communists. In another example, Muhyiddin claimed Jews and Christians had a covert agenda to proselytise and convert Muslims in Malaysia.



Twitter, 19 October 2022

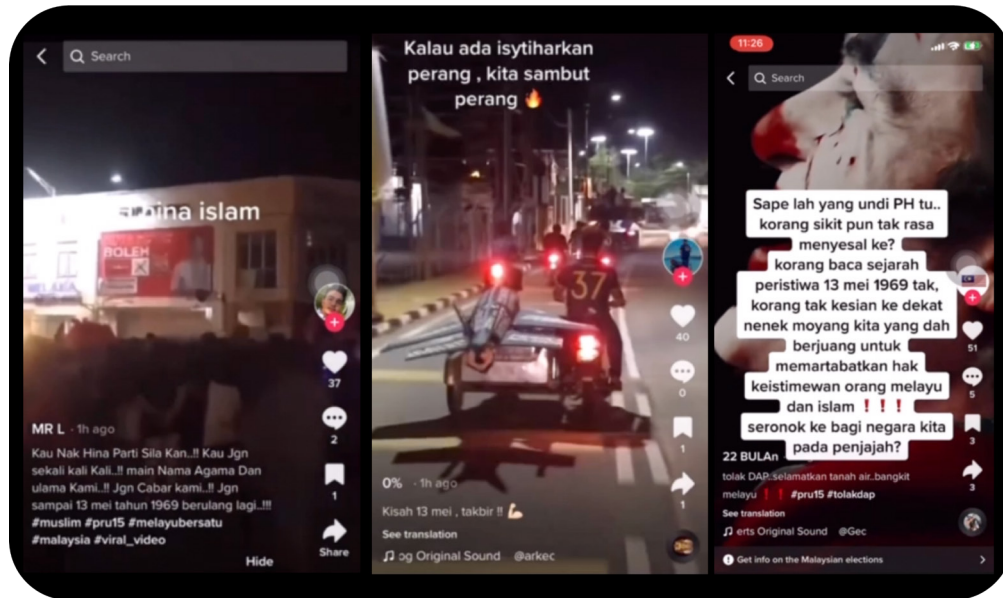
4.1.3 13 May 1969 and post-elections propaganda

The elections resulted in many firsts for Malaysia – the first hung parliament, the first time 18-year-olds could vote (and many voted for PN), and the first time PAS became the political party with the highest number of seats. In the days after the result was announced, it was



not clear who would be prime minister and what kind of coalition would form the government. This meant the Yang di-Pertuan Agong would play a bigger and unprecedented role in the formation of government.

The political uncertainty provided space for some actors to manufacture fear and intensify animosity. TikTok videos by young content creators that went viral threatened a possible recurrence of the historical race riots of 13 May 1969 should the DAP and PH return to power.



Tik Tok, 25 November 2022

Another prevailing narrative post-elections centred around disloyalty to the monarch and a push for a government that would uphold Malay sovereignty. This was mainly targeted at Muhyiddin and PN. These narratives were couched within the context of “derhaka” or treason. They also pushed for the monarch to intervene and break the political deadlock.

#### 4.1.4 Weaponization of the LGBTIQ issue

The targeting of LGBTIQ persons during GE15 was more politicised compared to past elections. LGBTIQ was rarely a standalone issue, as it was often coupled with claims of liberalism or anti-Islamic values. These narratives were often used against the PH coalition and its leaders. The coalition was seen as pro-LGBTIQ and therefore, anti-Islam. An exception was the use of the LGBTIQ narrative against PN’s Azmin Ali, but it was relatively muted compared to social media posts against PH, the DAP, or Anwar.

#### 4.1.5 Intersection between religion and other issues

As a growing phenomena in Malaysia we saw the visible shift in the manner in which religion was increasingly used to polarise an already divided society. The monitoring saw the trajectory of religion being used to stoke discordance and play into the insecurities among segments of Malay Muslims already experiencing socio economic hardship due to COVID 19 and economic slowdown. It became the critical fuse between protecting Malay rights and the need to defend Islam. The concept of “ketuanan Melayu” became synonymous with the concept of “ketuanan Melayu-Muslim”. Thus, the election campaign became the battle ground for the “Islamists” to push forth the ideology of needing to defend Islam from erosion by those branded as “secularists”, “deviants” or the “others”.

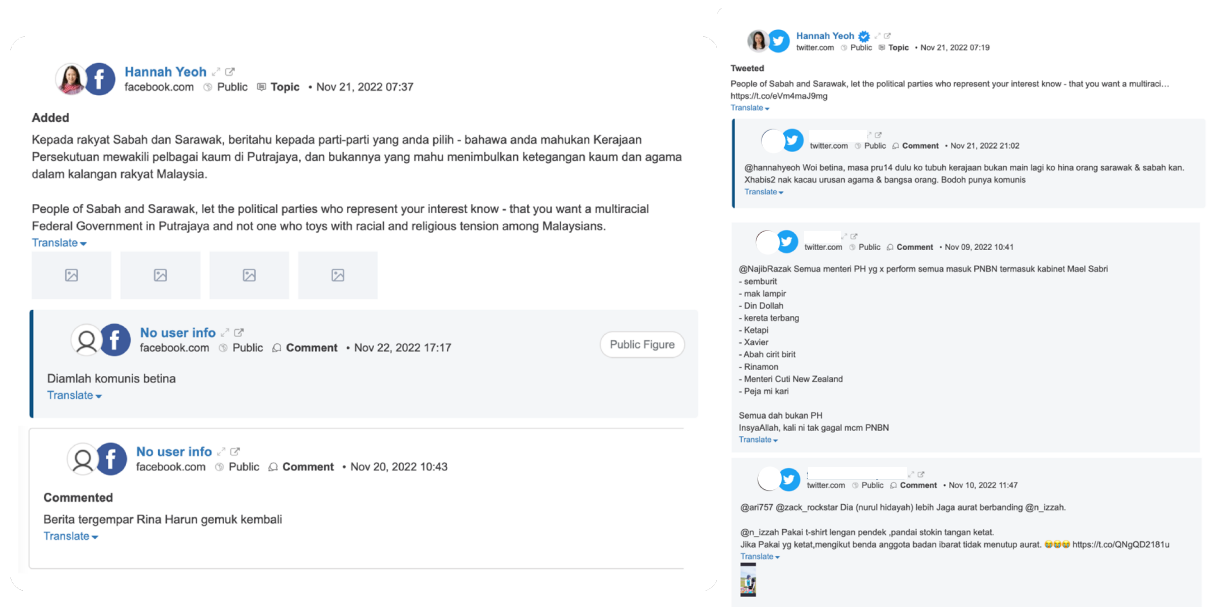
#### 4.1.6 Targeting of non-Malaysians

There was significant targeting of non-Malaysians on social media, with attacks on refugees and migrants reaching the highest severity level compared with the other issues that were monitored. While attacks against the LGBTIQ communities were weaponized as a political campaign tool, the same could not be said of the targeting of refugees and migrants. This reflects a wider problem of how Malaysia denigrates and dehumanises refugees and migrants, even more so since the COVID-19 pandemic.

The project decided to include refugees and migrants in this monitoring project because in the past, the state would sometimes round them up to divert public attention from policy issues. In the 2013 and 2018 elections, there were legitimate fears that large groups of migrant workers would be given Malaysian identity cards and bussed into polling stations. In 2022, this worry was virtually absent. Hence, the policing of refugees and migrants did not necessarily happen during this monitoring period. However, these groups were still consistently stigmatised and targeted, even if these actions were not linked to any political faction.

#### 4.1.7 Attacks against women politicians

Attacks against women politicians and candidates tended to revolve around body shaming and policing. An example was the focus on women to “tutup aurat” or to cover up. Women were also constantly reminded about the stereotypical roles of being mothers and homemakers. Worse were the derogatory labels used against women politicians such as “betina”. The term is used in the north and east coast of Peninsular Malaysia to mean female but could also denote a woman of dubious morals. Another label was “Mak Lampir” which refers to an evil, powerful witch in Javanese historical drama. These labels were frequently aimed at Rina Harun, Zuraida Kamaruddin, Hannah Yeoh and Nurul Izzah Anwar.



Facebook and Twitter, 9-22 November 2022

### 4.1.8 CIB

Almost one-third of the CIB posts observed during our monitoring was on Twitter. At 49.8%, almost half of CIB posts were on race and 38.9% on religion. The monitoring found it challenging to separate genuine social media accounts from fake ones in order to determine impact. This was because fake identity creators frequently employ sophisticated methods to make them appear authentic such as utilising profile images or imitating real users' behaviour.

## 4.2 Analysis by issues

### 4.2.1 Race

Among the five issues, race had the highest number of posts at 66,933 (refer to Table 4).

#### Actors

The highest number of posts about race originated from politicians and political parties. Ismail Sabri Yaakob, Anwar and PAS were the top three actors for all categories of actors and types of post under race.

The former prime minister, Ismail Sabri, had the most UGCs (6,229), which were mainly comments favourable to his administration's Keluarga Malaysia concept. They were low on the severity scale.

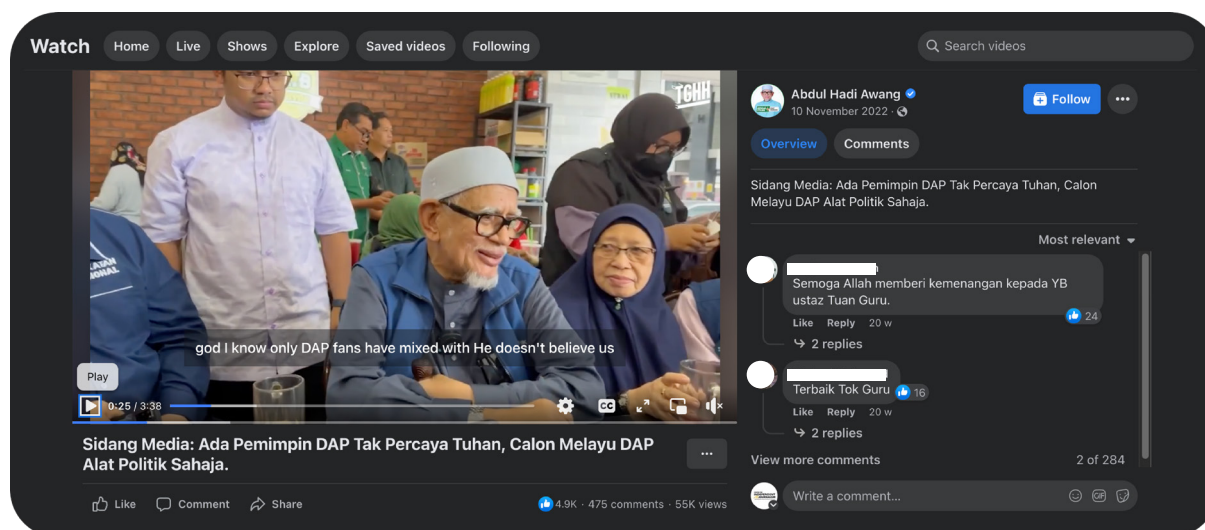
Despite not having the highest numbers with only 51 own posts, 1,803 UGCs and 1,926 mentions, Hadi was a huge amplifier on race.

#### Issues

##### a) Pushing anti-Chinese sentiments

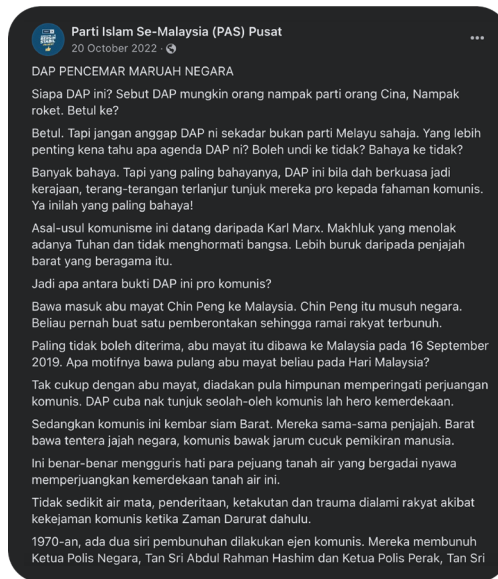
Hadi's posts, also reshared by the PAS party account and Harakah Daily, were negative against the DAP in the lead-up to the elections. These posts used terms such as "communist party", "liberals" and "atheists", and accused the DAP of corruption.

Hadi's TikTok post on 1 November called the DAP's fielding of Malay candidates a "dangerous and slick move".



Facebook, 10 November 2022





Facebook, 20 October 2022

Another example under this category was Najib’s posts on the chicken egg shortage, which he attributed to a cartel of Chinese companies.

### b) Malay supremacy

Up to 3,956 unique messages propagated “Ketuanan Melayu” (Malay supremacy). These posts went hand in hand with anti-Chinese rhetoric, and the concept of loyalty and obedience.

Riding on the huge success of the movie *Mat Kilau*, the story of Malay independence fighters was also used by other actors as a rallying call. Gerakan Pengundi Sedar’s post on Facebook on 26 October quoted Tok Gajah, a character in the film, who talked about defending race and religion (“jaga bangsa, agama”). An 18 November statement by PAS Pusat on Facebook, titled “Rintihan Buat Bangsa Melayuku” (Lament for my Malay race), criticised the PH manifesto. Quoting *Mat Kilau*, it called for the rise of Malays (“kebangkitan orang Melayu”).

### c) Threat of 13 May

The monitoring also found what seemed to be a well-coordinated and paid attempt on social media to generate anti-DAP and anti-Chinese fear and hatred, especially late on polling day. Inflammatory posts, that were mostly on TikTok, pushed an anti-DAP agenda while calling for a PN government.

These posts demonstrated trends<sup>25</sup> that include a rewriting of the history of 13 May, including that the DAP was the cause of the race riots. The posts then evolved into videos containing images of weapons and guns, with<sup>26</sup> messages warning the Malay majority to beware the DAP and PH. They also threatened a return of racial unrest at a time of deepened political tensions.

<sup>25</sup>[https://twitter.com/prvtbutnotso/status/1594656161789521921?s=46&t=VaFaiQBxN5Hra\\_HHWQu19g](https://twitter.com/prvtbutnotso/status/1594656161789521921?s=46&t=VaFaiQBxN5Hra_HHWQu19g)

- compilation of images of TikTokers spreading hate speech (general)

- Twitter account: RTU HARAPAN @prvtbutnotso

<sup>26</sup><https://twitter.com/bumilangit/status/1594666536152727552?s=46&t=OU1io8rD5hq4PGvV8Y7h4g>

- Compilation of images of TikTokers spreading hate speech and disinformation on 13 May 1969.

- Twitter account: Zulfadzli @bumilangit

<sup>26</sup> <https://twitter.com/bumilangit/status/1594666536152727552?s=46&t=OU1io8rD5hq4PGvV8Y7h4g>

There was significant engagement on this issue on TikTok. Additionally, these posts went viral across platforms, namely Twitter. Young content creators were mainly used to generate these video posts and netizens shared these contents aimed at creating fear, racial and religious polarisation. At times, they incited outright violence. There were allegations PN<sup>27</sup> had engaged professionals to coordinate these social media feeds.

#### d) Malaysia Agreement 1963 (MA63)

Data related to Sabah and Sarawak was insignificant, at only 1,052 unique messages for Sabah, and 1,674 for Sarawak. This low number could be partly due to our study’s limited scope which only tracked five politician actors from Sarawak and two from Sabah. For Sabah and Sarawak, the focus was not race or indigenous rights but the MA63 (70 posts with low severity level). Posts on these two states saw a hike after polling day, in relation to Anwar’s overtures to Gabungan Rakyat Sabah (GRS) to join the unity government.

It could also be hypothesised that politicians in East Malaysia were more focused on campaigning on the ground than on social media.

#### e) Insignificant targeting of Indians

While race had the highest number of posts, there was however very little data on Indians and Indian politics. There were only 616 unique messages, which were mostly posts alleging links to Sri Lanka’s Liberation Tigers of Tamil Eelam.

We had expected that the death of firefighter Muhammad Adib Mohd Kassim following a 2018 riot at a Hindu temple would be significant during the monitoring period. It wasn’t. It was only mentioned to question why the case had not been resolved.

#### Severity

Of all the issues monitored, race had the most number of posts with high severity levels. It had 53 Level 3 posts – the highest of all – and four Level 4 posts (refer to Table 7).

The 13 May TikTok videos which were uploaded around 19 November, immediately when the election result was announced, had huge amplification through sharing and repeats. Some of the young TikTok users who shared content on 13 May carried the “Paid Partnership” label, proof that their content was sponsored even though it remained unknown who the sponsors were.



TikTok, 25 November 2022

27 How Political Parties Used TikTok in the 2022 Malaysian General Election <https://www.rsis.edu.sg/rsis-publication/rsis/how-political-parties-used-tiktok-in-the-2022-malaysian-general-election/#.ZC-pDffmgWM>

Hate speech in response to PH’s posts were mainly triggered by remarks by the DAP’s Nga Kor Ming, particularly his statement that Malaysia would become a Taliban state if PAS ruled.

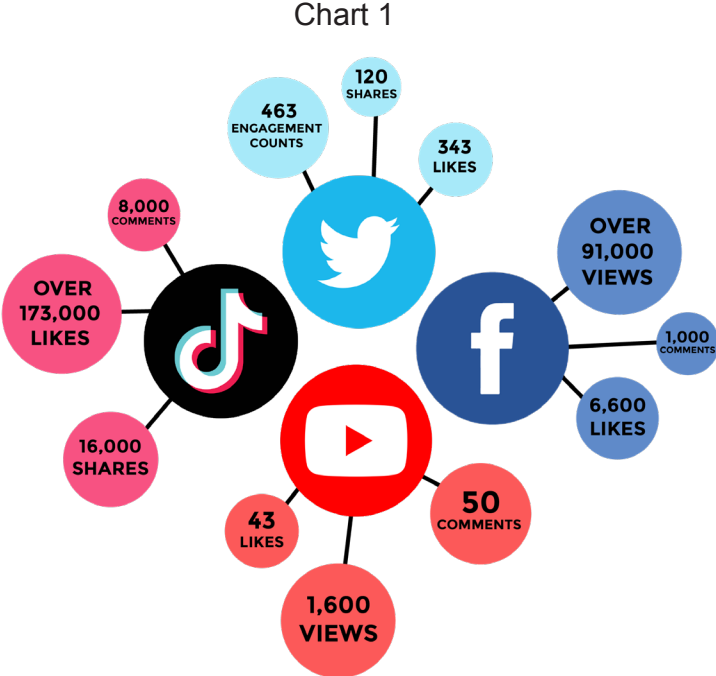
Table 7. Race: Severity Level and Type of Posts

SEVERITY LEVEL	TYPE OF POSTS			TOTAL POSTS
	Own Posts	UGCs	Mentions	
Level 1	988	34,560	18,461	54,009
Level 2	80	6,876	5,055	12,011
Level 3	2	36	15	53
Level 4	0	3	1	4
<b>Total Posts</b>	<b>1,070</b>	<b>41,475</b>	<b>23,532</b>	<b>66,077</b>

**Platform**

The majority of posts on race were found on Facebook and Twitter (refer to Table 8). For content on TikTok, race had the highest number of posts at 468.

Hadi’s TikTok video on DAP Malay candidates being used by the party generated the highest engagement across all platforms, at almost 2.5mil in engagement count. It had the following engagement count across the different platforms:



Apart from Hadi’s video, Mat Kilau also featured prominently on TikTok. It was used to represent the concept of Malay supremacy and the need to defend the Malay and Muslim position. This was then used by young creators to generate content on TikTok. Posts on Mat Kilau did not generate hate speech even though it was racist in nature.



Table 8. Race: Number of Posts and Platforms

	Facebook	TikTok	Twitter	YouTube	Total Posts
RACE	42,733	468	23,686	46	66,933

## CIB

Most of CIB activity was under race, with 1,458 accounts, or 49.83%, or almost half of all CIB observed by the monitoring project.

### 4.2.2 Religion

Religion had the second highest number of posts among the five issues, at 24,484 posts (refer to Table 4).

#### Actors

It was not surprising to see PAS, Hadi and the party organ Harakah dominated the narratives, whether in their own posts, UGCs or mentions.

Hadi alone had 121 own posts. Coming second under the politician category, Anwar had 20 own posts only but had a high amplification effect, with 1,535 UGCs. Hadi's posts appeared consistently across his personal accounts on all platforms, as well as on the PAS and Harakah Daily accounts. KOLs Firdaus and Isma's Aminuddin Yahaya also had high engagement on their own posts.

It was worth noting that Muda, a new political party, was third in terms of its own posts generating comments on religion. Out of 38 own posts, 36 were Level 1, and two were Level 2. Muda also featured prominently on the other issues of royalty, gender and LGBTIQ, and refugees and migrants. All their own posts were at Level 1.

Significant spikes in the number of messages and engagement on social media were seen in relation to Hadi's 22 October TikTok post about why PAS rejected PH ("Mengapa PAS tolak PH"). Spikes were also seen between 11 and 13 November with the amplification of videos by Jamal Abdillah, Zul Huzaimi, and Shukri Shaari from PAS Youth with their lobby for Muslim votes.

#### Issues

##### a) Defending Islam and sharia law

Islam was weaponized to fan Malay Muslim insecurities. The focus was often on the need for sharia, and in relation to that, Hukum Hudud<sup>28</sup> and the push for the controversial Rang Undang-Undang 355 (RUU355)<sup>29</sup> as key to upholding Islamic principles in Malaysia. These went hand in hand with attacks against the DAP being atheist and anti-Islam.

It should however be noted that there were no specific posts by Hadi, PAS or Harakah Daily leveraging RUU355 throughout the election campaign. Media coverage on RUU355 was also insignificant. The topic mainly came up in UGCs that originated from non-actors.

<sup>28</sup> Refers to Islamic penal law or Quranic punishments. It covers a range of crimes and contains punishments such as stoning for adultery, whipping for abortion and amputation for theft.

<sup>29</sup> A private member's bill brought by PAS president Abdul Hadi Awang in 2016 to amend the Shariah Courts (Criminal Jurisdiction) Act 1965 to give the law greater enforcement power.

The monitoring also found narratives portraying some politicians as defenders of Islam. Assertions were also made that not voting for PAS or PN would mean being un-Islamic or that the voter would likely “go to hell”.

### **b) Anti-LGBTIQ sentiment**

There was a clear intersection between religion and gender and LGBTIQ.

Religion was used to undermine and attack the rights of LGBTIQ persons. LGBTIQs were seen as deviant and going against Islamic precepts. The amplification of these posts was done by certain actors, specifically Harakah Daily and Merdeka Times and KOLs. (see more under gender and LGBTIQ below)

### **c) Attacks against Jews and Christians**

Towards the tail end of the campaign period, a video of Muhyiddin went viral. Muhyiddin alleged that a group of Jews and Christians, through proselytisation, were plotting for PH to rule Malaysia. However, engagement on this video was low and there were no spikes in the severity levels. In fact, social media users, including from Sabah and Sarawak, called him out for his disinformation.

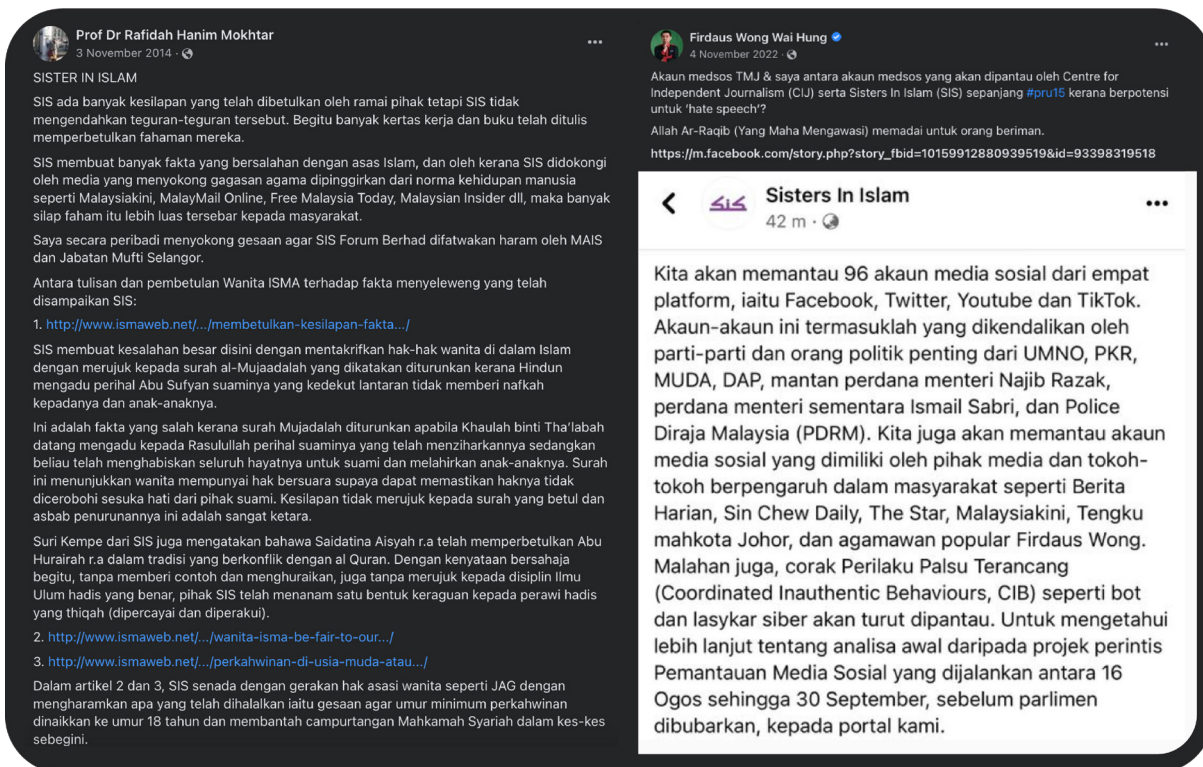
### **d) Public morality**

Women candidates were pressured to cover their aurat, including by wearing the headscarf, during the campaign period. Religion was often used as justification. There was also a spike in posts on polling day with claims that alcohol consumption and gambling would increase if the DAP came to power. This was used to assert that if the DAP was in government, it would fail to uphold Islamic values.

### **e) Use of KOLs to amplify messaging**

In several instances, celebrities or KOLs were used to amplify key messaging related to religion, including via cross-platform amplification. This was obvious with two viral videos. In one video, Malay singer Jamal Abdillah urged voters to reject the DAP, stoking fears that the azan (call for prayer) would be restricted and tahfiz schools closed should PH form government. In another video, actor Zul Huzaimy said at a PAS rally he harboured a wish to “slaughter infidels” (“kafir harbi”) and it was only the law which stopped him from doing so.

Two other actors we monitored, Firdaus and Dr Rafidah, did not directly refer to each other but they frequently amplified messages upholding Islamic principles and laws. They also singled out two CSOs – Justice for Sisters and Sisters in Islam – as being pro-LGBTIQ and anti-Islam.



Facebook, 3 & 4 November 2022

## f) Insignificant focus on child marriages and unilateral conversions to Islam

We found that no political faction focused on these issues as part of their political agenda.

### Severity

Many posts were at Level 1 (20,315) and Level 2 (3,545). Comments that were dehumanising or hostile were on the low side. There were eight Level 3 posts and two Level 4 posts (refer to Table 9).

The narrative for PAS was mainly positive. The party was referred to as “Islamic, clean and stable”. Their political rivals were either mocked or attacked. Nonetheless, there were pejorative terms used that referred to PAS as “Parti Pak Lebai<sup>30</sup>”.

UGC of PH was blunter. The coalition was deemed anti-Islam because it had the DAP as a component party and was painted as supporting apostates, atheists, LGBTIQs and liberals. Harakah Daily labelled PH a “communist” coalition to be rejected by voters.

<sup>30</sup> Shahrin Aizat Noorshahrizam, “Azmin came up with ‘Pak Lebai’ first, Zahid tells PAS”, 23 October 2022, <https://www.malaymail.com/news/malaysia/2022/10/23/azmin-came-up-with-pak-lebai-first-zahid-tells-pas/35200>

*Table 9. Religion: Severity Level and Type of Posts*

SEVERITY LEVEL	TYPE OF POSTS			
	Own Posts	UGCs	Mentions	TOTAL POSTS
Level 1	508	13,114	6,693	20,315
Level 2	70	2,084	1,391	3,545
Level 3	1	5	2	8
Level 4	0	1	1	2
<b>Total Posts</b>	<b>579</b>	<b>15,204</b>	<b>8,087</b>	<b>23,870</b>

## Platforms

The posts were primarily on Facebook and Twitter but TikTok content, at 445, was significant.

The coordinated and paid attempts to stoke anti-DAP and anti-Chinese fear and hatred through inflammatory posts were mostly on TikTok.

*Table 10. Posts on Religion and Platforms*

	Facebook	TikTok	Twitter	YouTube	Total Posts
RELIGION	15,284	445	8,710	45	<b>24,484</b>

## CIB

There were 1,138 CIB posts on religion, which was almost 40% of the total CIB posts. While the presence of CIB was apparent under religion, it did not impact the narrative. CIB in this category were insignificant in generating UGCs and did not play the role of amplifying hate speech, unlike the comments from regular social media users. When cross-referencing the data for CIB with political parties, the monitoring found that the highest was PAS (70), followed by Umno (13). These CIB were mainly pro-BN, judging by the coordinated attacks on PN and PH.

### 4.2.3 Royalty

Royalty ranked third in terms of the number of posts (14,320), compared to the other issues (refer to Table 4).

## Actors

Royalty as an issue had comparatively low numbers of own posts. The number was low even for the crown prince of Johor who led the KOL category. Like other issues, politicians dominated the top tiers of total posts.

One actor who stood out was Anwar, who had the most number at 3,490. Najib, despite being in prison, managed to attract the second highest number at 1,510.



Table 11: Royalty: Top 5 Actors

ACTOR	CATEGORY	TOTAL PROCESSED POSTS
Parti Keadilan Rakyat - Anwar Ibrahim	Politician	3,490
Umno - Najib Razak	Politician	1,510
Parti Islam Se-Malaysia	Political Party	1,430
Bersatu - Muhyiddin Yassin	Politician	1,313
Malaysian United Democratic Alliance	Political Party	1,230

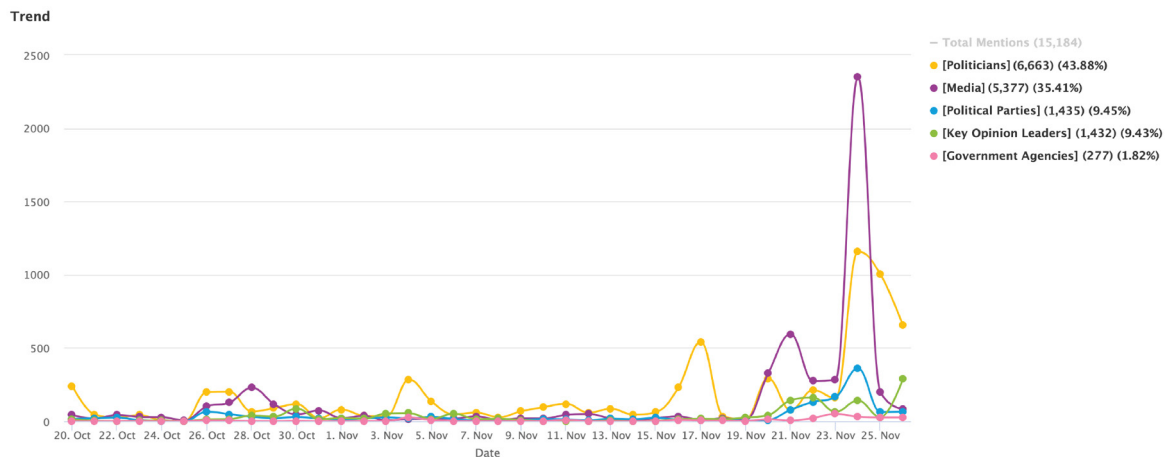
## Issues

### a) Royalty's key role amplified

There was a significant increase in posts related to royalty after the elections, especially leading up to the appointment of Anwar as prime minister on 24 November. These posts were made after the Agong's call for a unity government. They were primarily positive and supported the Agong's intervention in ensuring the nation's stability.

During this same period, there was also a spike in the number of UGCs critical of Muhyiddin for not joining the unity government. His decision was deemed "treason" against the Agong who had called for unity and stability.

Chart 2



## Severity

Most posts were at severity Levels 1 and 2, with only 19 at Level 3 (refer to Table 12). The 19 Level 3 posts were not against royalty but against other actors. For example, posts generated from news reports of Muhyiddin going to Istana Negara post-elections criticised him for being disrespectful of the Agong's call for parties to form a unity government. The target of hate speech under this sub-category was not the royal institution but Muhyiddin.

Table 12. Royalty: Severity Levels and Type of Posts

SEVERITY LEVEL	TYPE OF POSTS			
	Own Posts	UGCs	Mentions	TOTAL POSTS
Level 1	596	9,204	3,196	12,996
Level 2	12	630	289	931
Level 3	0	19	0	19
Level 4	0	0	0	0
<b>Total Posts</b>	<b>608</b>	<b>9,853</b>	<b>3,485</b>	<b>13,946</b>

### Platform

Most of the posts were on Facebook and Twitter.

When filtering for the keyword “derhaka” (treason), Twitter had the highest engagement with 396 messages.

Table 13. Royalty: Platforms

	Facebook	TikTok	Twitter	YouTube	Total Posts
ROYALTY	10,078	81	4,113	48	14,320

### CIB

CIB for royalty was low, at less than 5%, or 134 posts.

### 4.2.4 Gender and LGBTIQ

Of all five issues, gender and LGBTIQ had the second lowest number of social media posts at 7,596 (refer to Table 4).

### Actors

For a new media organisation, Merdeka Times featured prominently, with the highest number of its own posts (64), ahead of political parties, politicians, and established media . PAS was second, at 40 posts, followed by Merdeka Times’ strategist and KOL Firdaus (39). Both Firdaus and PAS stood out not only in their own posts but also in UGCs and mentions by social media users.

Table 14: Gender and LGBTIQ: Top 5 Actors

ACTOR	CATEGORY	TOTAL PROCESSED POSTS
PKR - Anwar Ibrahim	Politician	1,135
Parti Islam Se-Malaysia	Political Party	1,000
Bersatu - Azmin Ali	Politician	766
Firdaus Wong Wai Hung	Key Opinion Leader	719
Democratic Action Party	Political Party	694

There was a surge of posts specifically on LGBTIQ throughout the monitoring period. One example was the 37 posts during the entire monitoring period that called Azmin “semburit”, a derogatory term to describe gay anal intercourse. Another spike on 15 November was connected to a number of posts referring to Syed Saddiq as “pondan”, a derogatory term to describe effeminate men.

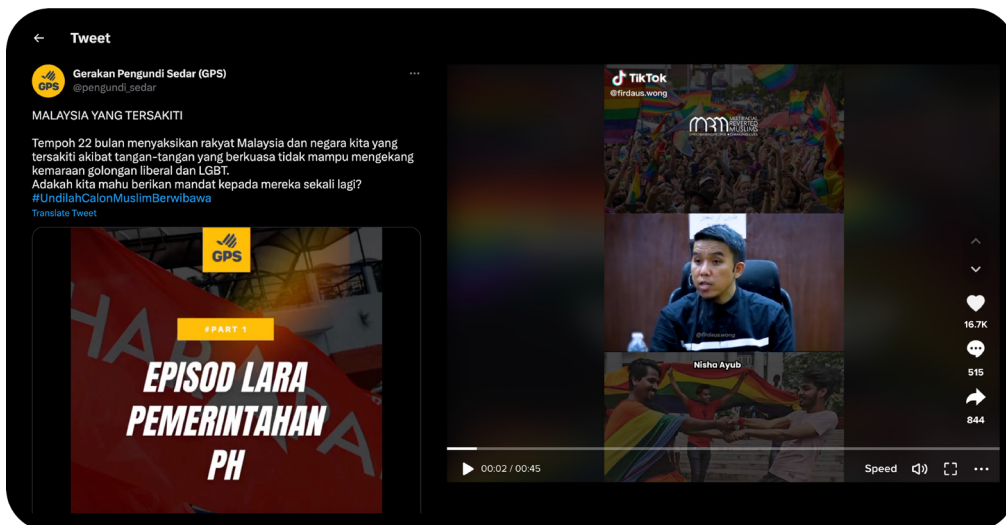
## Issues

### a) Targeting LGBTIQ persons

Misogynistic opinions and arguments on the basis of public morality and Islamic values were weaponized to incite hate against the LGBTIQ community.

Comments on LGBTIQ were mostly aimed at discrediting politicians, with PKR’s Anwar and Bersatu’s Azmin being the key targets and featuring prominently in UGCs and mentions. Posts claimed that Anwar, PKR and PH supported LGBTIQ and would include an LGBTIQ agenda if voted into power.

Together with Gabungan Pengundi Sedar, Firdaus and Dr Rafidah were effective amplifiers, retweeting and sharing each other’s posts. Firdaus and Dr Rafidah cited an Islamic legal perspective to justify their opposition to LGBTIQ having equal rights. They claimed LGBTIQ sexuality was a mental health issue and utilised articles from conservative Western media to support their point of view.



Twitter, 7 November 2022 and TikTok 2 November 2022

Posts on LGBTIQ specifically targeted gay and transgender persons, with almost no specific content targeting lesbians.

On 29 October, police and the Federal Territory Islamic Religious Department (Jawi) raided a private drag queen Halloween party in Kuala Lumpur. Subsequently, individuals and organisations representing LGBTIQ rights were targeted in a seemingly coordinated manner across various platforms such as Facebook, Twitter and TikTok. Activist Numan Afifi quickly became a target online. Comments ranged from congratulating the authorities for the raid and related arrests, to calling for the elimination of the LGBTIQ community, and falsely claiming that the community spread a viral disease.



Facebook, 25 November & Twitter 31 October 2022

Dr Rafidah's post, thanking the authorities for their action, triggered more hateful comments. Politicians surprisingly did not comment on news of the raid. Only the DAP's Charles Santiago and the PSM Youth Wing released press statements decrying the authorities' action.



Facebook, 30 October 2022



## b) Gender stereotypes

Some posts policed and objectified women candidates and public figures, subjecting them to body shaming and derogatory remarks. For example, independent candidate and social media influencer Nur Fathiah Syazwana Shaharuddin (known by her handle of Cleopatra) and Najib's daughter Nooryana were instructed to "tutup aurat" or cover up with the headscarf.

Bersatu's Rina and PKR's Nurul Izzah were referred to as "gemuk" (fat), DAP's Yeoh as "betina gharik" (wanton woman), and DAP's Young Syefura as "betina Melayu pengkhianat" (female Malay traitor).

The focus on how a woman looks draws attention away from her skills, expertise or actions, and can influence voters to perceive women as less credible or equipped for public office. These prevailing sexist and gender-based hate narratives undermine the role of women in politics, and limit women's right to political participation.



Twitter, 13 November 2022

Kedah Menteri Besar Muhammad Sanusi Md Nor, in a sexist TikTok video, also belittled and undermined women candidates in Sik by saying they were not competent or capable leaders. The media then reported Azmin's support of Sanusi's remarks, hence amplifying the sexism. The use of "betina" is not considered derogatory in Kedah, but the tone and content of Sanusi's remarks were discriminatory and disparaging of women.

Objectification and body shaming of male politicians also happened but to a much lesser extent. For example, Tengku Adnan Tengku Mansor, popularly known as Ku Nan ("gemuk" or fat) and Khairy Jamaluddin ("gelap" or dark skinned). Posts referring to Syed Saddiq as "pondan" undermined his credibility as a political leader.



Facebook, 29 October 2022

**c) Gender and citizenship**

The monitoring found specific references to the right of Malaysian women to confer citizenship to children born outside of Malaysia to non-Malaysian fathers, as one of the election topics.

**Severity**

Gender and LGBTIQ had the least number of posts with severity Levels 3 and 4. Most were at Level 1 (4,089) and Level 2 (4,212) (refer to Table 15).

The Level 1 and 2 posts called for use of violence, specifically against LGBTIQ persons, including calling for “sula” (impaling with a sharp object), “kasi pancing” and “kasi libas” (find them, whip them), amongst others.

*Table 15. Gender/ LGBT Issues - Severity Level and Type of Posts*

SEVERITY LEVEL	TYPE OF POSTS			
	Own Posts	UGCs	Mentions	TOTAL POSTS
Level 1	215	2,752	1,122	4,089
Level 2	106	2,592	1,514	4,212
Level 3	0	4	2	6
Level 4	0	0	0	0
<b>Total Posts</b>	<b>321</b>	<b>5,348</b>	<b>2,638</b>	<b>8,307</b>

**Platform**

Two thirds of all posts on this issue were on Facebook, followed by Twitter at 31.9% (refer to Table 16).

*Table 16. Gender/LGBTIQ Issue and Platforms*

	Facebook	TikTok	Twitter	YouTube	Total Posts
<b>GENDER AND LGBTIQ</b>	<b>5,038</b>	<b>121</b>	<b>2,427</b>	<b>10</b>	<b>7,596</b>

**CIB**

There were 81, or 27.68%, CIB posts identified under gender and LGBTIQ, which is just above a quarter. Most of them were on Anwar and Azmin’s purported sex videos and photographs. They were shared frequently by CIB.

**4.2.5 Refugees and migrants**

Of the five issues, refugees and migrants had the least number of social media posts at 3,819 (refer to Table 4).

## Actors

One actor which stood out was the Immigration Department. It got the highest number of UGCs on the issue at 785. The department had 63 own posts, the highest among all the actors. This was understandable, given that the department is the official government channel on the issue of refugees and migrants.

Another notable actor was PAS, which was second in UGCs (320) and own posts (13). PAS had the most posts compared to all political parties.

*Table 17. Refugees and Migrants: Top 5 Actors*

ACTOR	CATEGORY	TOTAL PROCESSED POSTS
Immigration Department of Malaysia	Government	958
Parti Islam Se-Malaysia	Political Party	500
Friends of Immigration	Key Opinion Leader	292
Berita Harian online	Media	276
Malaysian United Democratic Alliance	Political Party	248

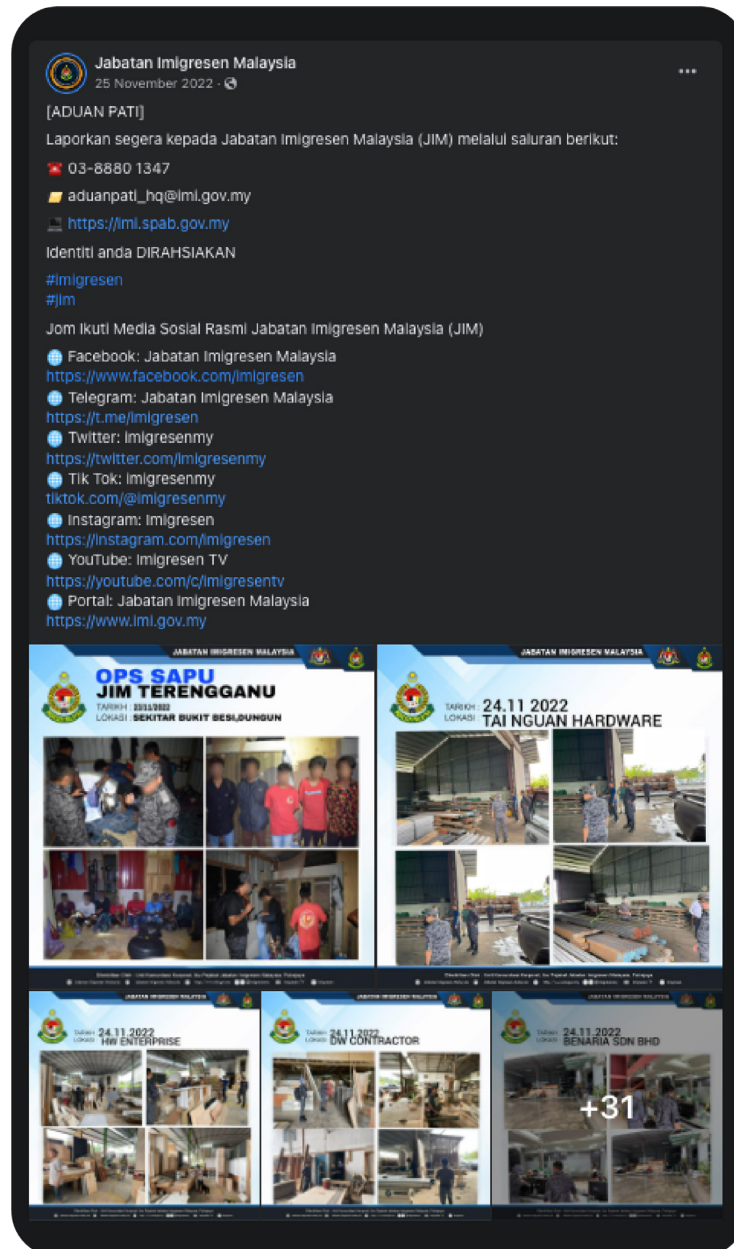
## Issues

### a) Xenophobia and targeting of refugees and migrants

Negative portrayals of both communities have been consistent over the years, including in the run-up to the elections. Most posts sow hatred by claiming these communities threaten national security and Malaysians' job security.

Monitoring has found that such conversations on social media lead to heightened hate speech, including instances of doxing and incitement of violence. However, our monitoring did not find clear attempts to use this issue as part of any political agenda.

We found that the spike in hate speech against refugees and migrants was often triggered by repeated Immigration Department posts encouraging the public to report undocumented migrant workers via its hotline. The department's posts and raids resulted in hateful comments, including calls for "extermination", and doxing where migrant workers' contact details were revealed on social media.



Facebook, 25 November 2022

We recorded an increase in online attacks on 28 October, and on 5, 12 and 24 November, following either the department's social media posts or its Ops Dandan, Ops Kedah and Ops Gebeng raids.

Xenophobic and anti-migrant sentiment was also evident in an interview with DAP's Ganabatirau Veraman when he spoke about migrant workers in Meru Market.

### b) Statelessness

The number of posts about stateless children in Sabah, and about the BN's purported issuance of identification documents to non-Malaysians to garner votes, was insignificant during our monitoring.

We also found that Dayak Daily and Borneo Post did not amplify the issue of statelessness in Sabah.

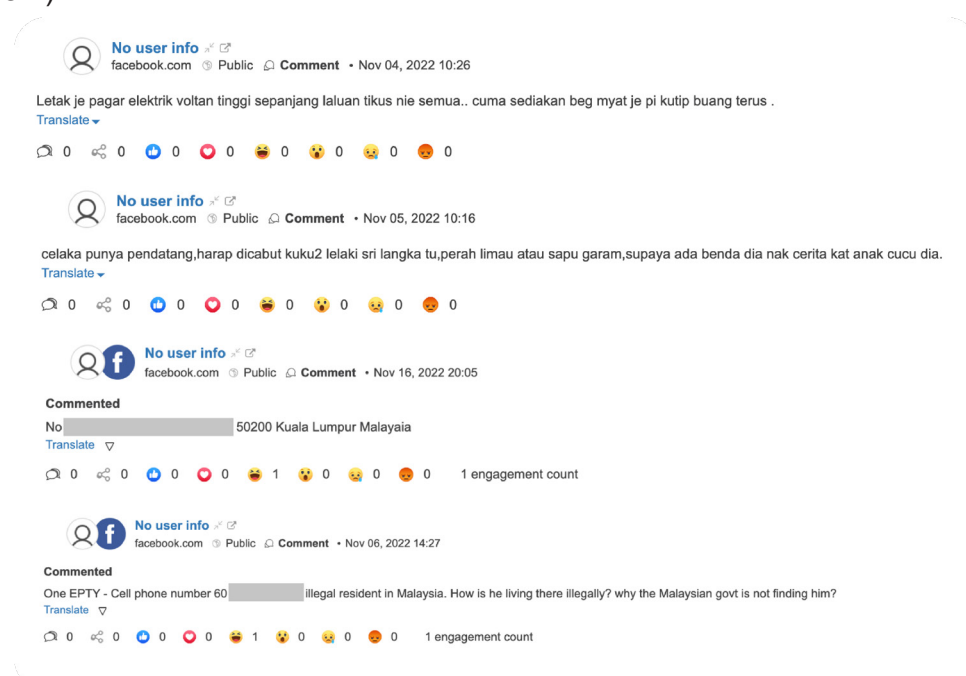


## Severity

At 3,819, the total number of posts on refugees and migrants was low compared to the other issues (refer to Table 4). However, they generated the highest number of Level 4 posts with all 28 of these posts being UGCs (refer to Table 18).

The Immigration Department's own posts on Facebook were only at Levels 1 and 3, and hence did not constitute hate speech. However, the page admin's own posts encouraged the public to submit complaints and details on the whereabouts of refugees and migrant workers. In response, some UGCs shared the phone numbers and locations of migrants which they suspected to be without documents. These posts were registered as doxing and reported as such under the respective platforms.

Some examples of UGCs tagged as Level 4 contained explicit calls for physical harm, damage or death: "letak je pagar eletrik voltan" (just put up high-voltage electric fences); "cabut kuku" (pull out their nails); "tembak sampai mati" (shoot to kill); and "tembak je" (just shoot them).



Facebook, 4 - 16 November 2022

Table 18. Refugees and Migrants: Severity Levels and Type of Posts

SEVERITY LEVEL	TYPE OF POSTS			
	Own Posts	UGCs	Mentions	TOTAL POSTS
Level 1	122	2,052	507	2,681
Level 2	12	578	237	827
Level 3	1	17	6	24
Level 4	0	28	0	28
<b>Total Posts</b>	<b>135</b>	<b>2,675</b>	<b>750</b>	<b>3,560</b>

## Platform

Most posts on this issue were on Facebook at 2,969 posts, followed by Twitter 770 (refer to Table 19). The majority of Levels 3 and 4 posts were UGCs on the Immigration Department's Facebook page, the Friends of Immigration's community page, and the Berita Harian and Sinar Harian's pages.

Table 19. Refugees and Migrants: Number of Posts and Platforms

	Facebook	TikTok	Twitter	YouTube	Total Posts
REFUGEES AND MIGRANTS	2,969	73	770	7	3,819

## CIB

CIB posts totalled 115, or 3.93%, the lowest count among all the issues.

### 4.3 CIB

Our monitoring identified 2,623 messages as being posted by CIB accounts. These posts made up 2.63% of the total unique messages of 99,563. Most CIB activity was found on Twitter (1,642 posts), and most CIB activity was focused on race (1,458 posts) and religion (1,138 posts).

Our findings reaffirmed the conclusions of other research about attempts to control public narrative during elections in Malaysia. According to the Oxford Internet Institute (OII)'s Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation, both bots and cybertroopers were used to influence the information landscape in Malaysia's elections. Published in January 2021, the OII Report's case study on Malaysia<sup>31</sup> found CIB use was reported as early as the 2008 elections. In another study, a machine scan by the Digital Forensic Research Lab identified 17,000 bots targeting PH on Twitter ahead of the 2018 general election.<sup>32</sup>

Although the number of posts by CIB accounts is less than 10% of the total number of unique messages in our monitoring project, 2,623 posts is still a significant number considering we only monitored five issues.

The identification process of CIB accounts for this project was based on several rules<sup>33</sup>. However, it was difficult to distinguish between bots and cybertroopers on Facebook, TikTok and YouTube. The monitors used Botometer<sup>34</sup> to determine accounts with bot behaviour on Twitter. Because of this limitation, we defined bots and cybertrooper behaviour as CIB activity. The monitors also recognised CIB by noticing strange or long account usernames. Some accounts used politicians' names as usernames but further inspection identified these accounts as CIB.

31 Bradshaw, S., Campbell-Smith, U., Henle, A., Perini, A., Shalev, S., Bailey, H. & Howard, P.N. (2021) Country Case Studies Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation, p. 250 - 258. [https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/03/Case-Studies\\_FINAL.pdf](https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/03/Case-Studies_FINAL.pdf)

32 #BotSpot: Bots Target Malaysian Elections (2018). DFRLab. Retrieved from <https://medium.com/dfrlab/botspot-bots-target-malaysian-elections-785a3c25645b>

Accessed on 6 April 2023

33 Refer to 'Rules in determining inauthentic accounts' in Chapter 2 Monitoring Framework, p. 14 of this report.

34 Botometer investigates a Twitter account's activity and gives it a score. Higher scores indicate more bot-like activity. Link: <https://botometer.osome.iu.edu/>

### 4.3.1 CIB, actors and issues

Most CIB accounts mentioned politicians such as Rafizi, Anwar, Ismail Sabri and Muhyiddin in a total of 2,002 posts and tweets. Key accounts that were identified as CIB were Ayahanda MD, BossBM, TeamBossKu, Kuli Keng and KroniRakyat. Most of these accounts were activated months and even years prior to GE15. But they became active during the election campaign period and slightly after polling ended.

These accounts tagged politicians, political parties, and each other under the same tweets and posts. Discussion threads on religion and race indicated that certain accounts, such as AyahandaMD, TeamBossku and Bossbm, interacted with each other.

**AyahandaMD** • Oct 28, 2022 16:22

Tweeted

Patutlah jutaan warga cina hilang tdk dpt diksn ketika PH menerajui negara rupanya banyak tawaran spt ini dibuat oleh pemimpin PH kpd warga asing. Tindakan PH ini bkn sahaja merugikan org melayu tetapi kpd semua anak kelahiran m̄y tdk mengira kaum! @umnoonline @Alamurni @zlk69 <https://t.co/jwZBWcmS9F>

**Bossbm** • Oct 28, 2022 16:24

RT @Ayahanda\_MD: Patutlah jutaan warga cina hilang tdk dpt diksn ketika PH menerajui negara rupanya banyak tawaran spt ini dibuat oleh pemi...

**Bossbm** • Oct 28, 2022 16:24

@Ayahanda\_MD @umnoonline @Alamurni @zlk69 @bossku\_team @fikhwanBait @RameshRaoAKS @MakkalSaktiOrg @hamimahothman4 @romonICASTRO @NikYahya2 @NazzirHussain @hanifhanif88 Dah gila apa? Koming Ni..  
Pondan kaki block..

Twitter, 28 October 2022

**TeamBossku #SokongBN Dulu, Kini & Selamanya.** • Nov 05, 2022 07:20

RT @nverquestionme: @bossku\_team @miqaelridzuan @SIFIRMUDAH @Ayahanda\_MD @MuhyiddinYassin @umnoonline @Alamurni @Bossbm17 @fikhwanBait @zlk69 @RameshRaoAKS @BNrasmi @DrZahidHamidi @MakkalSaktiOrg @hamimahothman4 @romonICASTRO @Vandeeta3 @fatthebee @CikSueCinta @Shaffielsmail Klu nk rjuk smula kne tggu slps PRU15  
Rjuk awai² cnfrm trbongkar pekong PAS lg  
DAP ni party trbuang  
Dok mrempat sna sni for collaboration  
Tp bila dpt kuasa, lu sapa wa sapa  
Last² suma kne ikut kata DAP  
Maka mingkup la Melayu skali lg  
MELAYU MUDAH LUPA !

**TeamBossku #SokongBN Dulu, Kini & Selamanya.** • Nov 05, 2022 07:22

@nverquestionme @miqaelridzuan @SIFIRMUDAH @Ayahanda\_MD @MuhyiddinYassin @umnoonline @Alamurni @Bossbm17 @fikhwanBait @zlk69 @RameshRaoAKS @BNrasmi @DrZahidHamidi @MakkalSaktiOrg @hamimahothman4 @romonICASTRO @Vandeeta3 @fatthebee @CikSueCinta @Shaffielsmail Saya dah cukup masak dgn perangai cina DAP ini. Depa cuba berbaik dgn kita bila nak minta sesuatu. Depa hidup bermuka2!!

**Bossbm** • Nov 05, 2022 08:12

RT @nverquestionme: @bossku\_team @miqaelridzuan @SIFIRMUDAH @Ayahanda\_MD @MuhyiddinYassin @umnoonline @Alamurni @Bossbm17 @fikhwanBait @zlk...

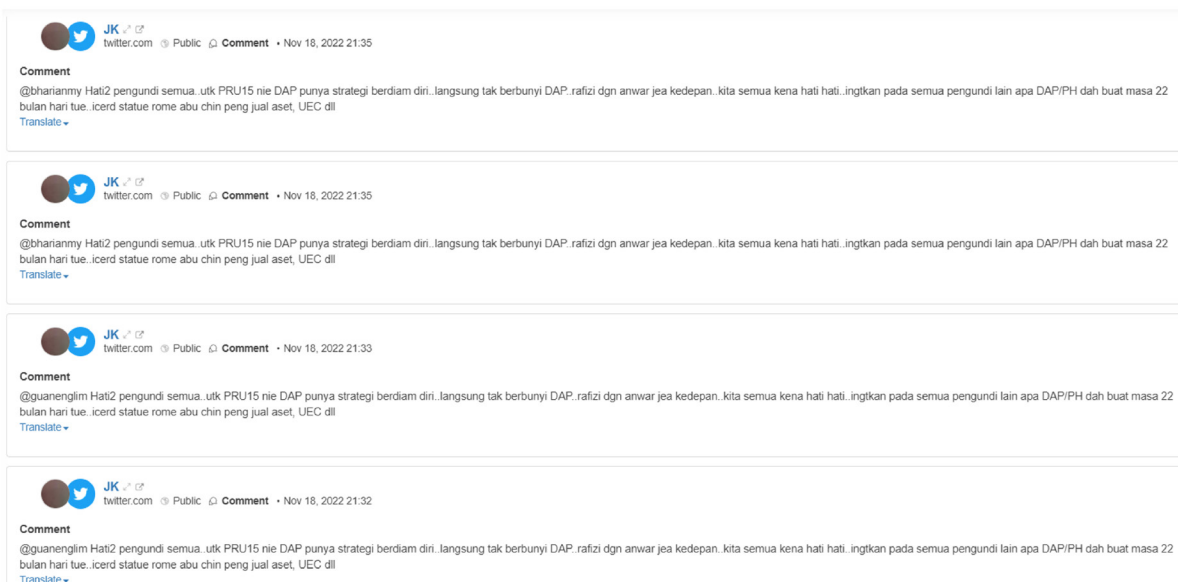
Twitter, 5 November 2022

CIB activity was also identified under posts made by the media such as Berita Harian and Malaysiakini.

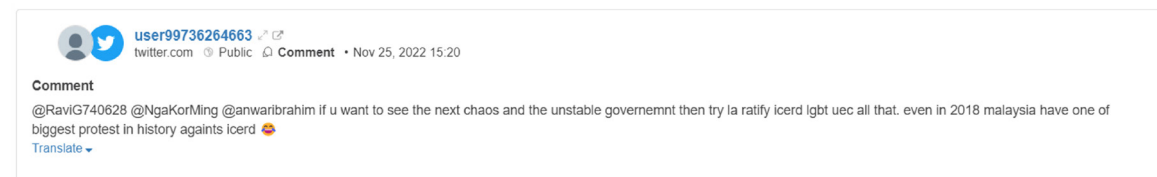


Facebook, 22 October 2022<sup>35</sup>

Another example of CIB posts or comments that stoked racial or religious discontent was about the Unified Examination Certificate.



Twitter, 18 November 2022



Twitter, 25 November 2022

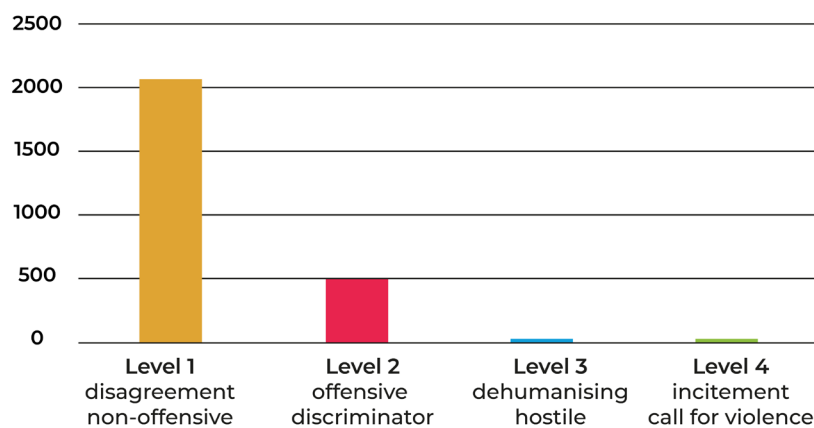
### 4.3.2 CIB and severity levels

At 2,090 posts, most of the CIB posts were at Level 1. These posts were mainly disagreements and contained non-offensive language. This was followed by 507 Level 2 posts which were offensive and discriminatory.

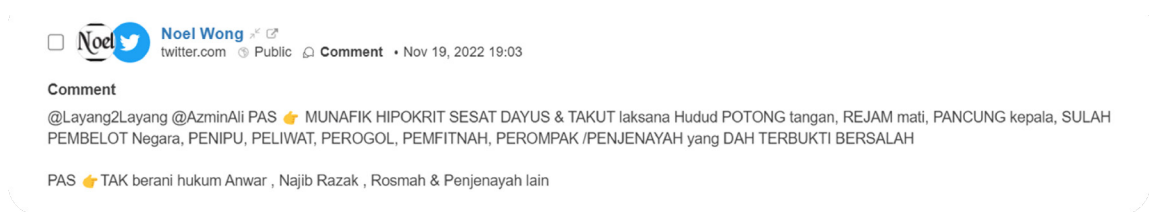
<sup>35</sup> The first post was a comment found under Berita Harian Online Facebook post on Anwar contesting in Tambun. This comment was considered Level 3. The second post was a comment found under Malaysiakini Facebook post about Perak PKR voters. Both posts were posted by the same CIB account on the same day with similar language. Both comments were tagged as Level 3.



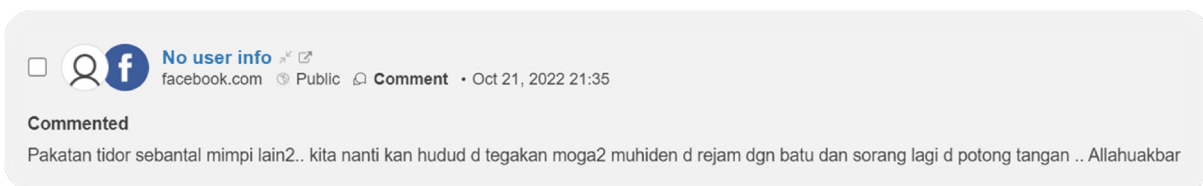
Chart 3: Severity level of CIB posts



There were more Level 4 posts (19), with language calling for violence and incitement, than Level 3 posts (7) that contained dehumanising or hostile language.



Facebook, 19 November 2022<sup>36</sup>



Facebook, 21 November 2022<sup>37</sup>

We found that the highest engagement between CIB and politicians was on religion. The monitoring also found that several CIB accounts interacted with each other. They also tagged politicians and mentioned political parties. The small but significant number of CIB posts on the issues we monitored indicate that not much CIB activity is needed to stoke discussions or propagate hateful messages online around these issues. We infer from this that it doesn't take much to get netizens to share or post hateful narratives on these issues.

#### 4.4 Analysis of social media platforms

The bulk of posts were on Facebook (63,643), followed by Twitter (34,782), TikTok (993) and YouTube (145) (refer to Table 20).

We experienced challenges in the automated scraping of data on TikTok due to limited access to its API. To enrich the analysis, the TikTok dataset was augmented by manual scraping. This decision proved useful when TikTok emerged as a key platform for hateful videos post-election day.

<sup>36</sup> Comment posted under Azmin's tweet about winning the elections. This comment was categorised as Level 4.

<sup>37</sup> Comment posted under Hadi's Facebook post about Himpunan Penyatuan Ummah. This comment was categorised as Level 4.

We could not provide a quantitative breakdown of data for TikTok as we did for the other platforms because of the manual scraping. Hence, in order to align the analysis with this research project’s objectives, we focused on TikTok posts tagged at Levels 3 and 4 only. In the following section, TikTok posts gathered from automatic scraping will be labelled as such to distinguish it from posts obtained from manual scraping.

*Table 20: Platform and Posts*

PLATFORM	POSTS
Facebook	63,643
TikTok (automated scraping)	993
Twitter	34,782
YouTube	145
<b>Total</b>	<b>99,563</b>

#### 4.4.1 Actors

For every category of actor monitored, there were more posts on Facebook compared to other social media platforms. This was followed by Twitter, TikTok and then YouTube, even though TikTok has been around less than half the time of YouTube’s existence.

Syed Saddiq was the actor with the highest engagement on TikTok, followed by Rafizi. Lim, even for his own posts which did not attack anyone, had the most negative comments, with words such as “komunis” (communist) and “pendatang” (immigrant) levelled at him. TikTok content produced for Hadi, who has almost 170,000 followers and 1.2 million likes, did not appear much on the automated system nor the monitors’ For You Page (for manual scraping), despite his popularity.

*Table 21: Category of Actor and Platform*

CATEGORY OF ACTOR	PLATFORM				
	Facebook	TikTok (automated scraping)	Twitter	YouTube	Total
Politicians	39,000	794	25,306	74	65,174
Political Parties	22,970	206	15,714	57	35,947
Media	21,454	0	7,542	125	29,121
Key Opinion Leaders	4,902	152	2,098	6	7,158
Government Agencies	3,001	62	1,533	6	4,602
<b>Total</b>	<b>91,327</b>	<b>1,214</b>	<b>52,193</b>	<b>268</b>	<b>145,002</b>

#### 4.4.2 Issues

For all the issues we monitored, around two thirds of total posts were on Facebook, followed by Twitter, TikTok and YouTube (refer to Table 22).

The 13 May videos which were posted on TikTok towards the end of 19 November tapered off after 24 November. As described above, some videos were labelled “paid partnership”, suggesting they had been commissioned. Others had the hashtag #ditaja (sponsored), although this was not necessarily evidence of sponsorship.

LGBTIQ issues were mostly brought up by Firdaus, who rarely talked about politics on TikTok. With the minimum voting age lowered to 18, TikTok may have been the perfect platform to reach out to younger voters who form the bulk of TikTok users.

Table 22: Issue and Platform

ISSUES	PLATFORM				
	Facebook	TikTok <small>(automated scraping)</small>	Twitter	YouTube	Total
Gender and LGBTIQ	5,038	121	2,427	10	7,596
Race	42,733	468	23,686	46	66,933
Refugees and migrants	2,969	73	770	7	3,819
Religion	15,284	445	8,710	45	24,484
Royalty	10,078	81	4,113	48	14,320
<b>Total</b>	<b>76,102</b>	<b>1,188</b>	<b>39,706</b>	<b>156</b>	<b>117,152</b>

#### 4.4.3 Severity

The bulk of posts on all platforms was at Level 1 (81,763) and Level 2 (17,656). There were 105 Level 3 posts (84 on Facebook and 21 on Twitter), and 39 Level 4 posts (36 on Facebook and three on Twitter) (refer to Table 23).

Sources of hate speech on TikTok were found through manual scraping. There were 35 Level 3 posts and four Level 4 posts (refer to Table 24). The actors responsible for these posts were KOLs Firdaus, Syahir Norhisham, and Superman Hew (linked to the DAP), and politician Sanusi. Other sources included media Harakah Daily, and young adults on TikTok who urged viewers to support PN, and not vote for the DAP.

The most prominent narrative on TikTok from our manual scraping was the demonisation of the DAP. This was done through racial incitement with the 13 May videos. It was also done by generating anti-DAP and anti-communist sentiments, labelling PH supporters as anti-Malay, and asserting pro-Malay nationalist views, including calling non-Malays “pendatang” (immigrants). Comments on these videos called for Malays to “bangkit” (rise up) and fight those who do not support the Malay agenda. Young adults and teenagers became amplifiers of this hate speech on TikTok. In videos posted after election day, a handful of TikTok users flashed weapons and called on other Malays to stand up against the non-Malays whom they painted as disloyal to Malaysia.

Table 23: Severity and Platform

SEVERITY LEVEL	PLATFORM				
	Facebook	TikTok (automated scraping)	Twitter	YouTube	Total
Level 1	53,037	776	27,814	136	81,763
Level 2	10,486	217	6,944	9	17,656
Level 3	84	0	21	0	105
Level 4	36	0	3	0	39
<b>Total</b>	<b>63,643</b>	<b>993</b>	<b>34,782</b>	<b>145</b>	<b>99,563</b>

Table 24: Severity and TikTok (manual scraping)

SEVERITY LEVEL FOR TIKTOK (manual scraping)	POSTS
Level 3	35
Level 4	4
<b>Total</b>	<b>39</b>



# SOCIAL MEDIA MONITORING OF MALAYSIA'S 15TH GENERAL ELECTIONS



## 5 | RECOMMENDATIONS

## 5. Recommendations

The project team has these recommendations for various stakeholders.

### 1. State

a) Set up a multistakeholder Independent Commission to address the escalation of harmful narratives and hate speech. The Independent Commission should be tasked to:

- Review the root causes and drivers of harmful narratives and hate speech, especially in the context of race and religion;
- Consult with civil society organisations to better understand the nature of hate speech and its likelihood of harm and how to proportionately respond to hate speech; and
- Develop concrete actions and recommendations, across different actors and platforms, and in line with international standards, specifically the Rabat Plan of Action.

b) Initiate more comprehensive media and digital literacy programmes to educate students and public on the dangers of hate speech and promote equality, non-discrimination, diversity and inclusivity.

c) Policymakers must ensure that any measures adopted to combat hate speech do not unduly impair freedom of expression and speech. Any restrictions to freedom of expression and speech must meet international human rights standards of legitimacy, necessity and proportionality.

### 2. Politicians/political parties

a) Politicians, along with other public figures, have a vital role to play in this process. Politicians have both a political obligation and an ethical responsibility to refrain from using hate speech and stigmatising language and to condemn promptly and unequivocally its use by others.

b) Politicians and political parties are to develop and adopt self-regulatory instruments such as guidelines or policies, which are particularly effective and more likely to be respected due to their voluntary nature.

### 3. Election Commission of Malaysia (EC)

a) Adopt clear guidelines to prevent and ensure political parties and candidates refrain from hate speech during campaigns and impose necessary penalties for such action.

b) Relatedly, push for amendments to Section 4A (1) and Section 9 of the Elections Offences Act to explicitly address hate speech during elections.

### 4. Social media platforms

a) Ensure its community standards and practices are centred on human rights and corporate accountability in line with the United Nations Guiding Principles on Business and Human Rights and other international human rights standards.

b) Put resources into timely detecting and responding to user complaints about removing harmful messages on their platforms. Companies like Meta, TikTok, Twitter and Google can choose to divert additional resources into content moderation during pivotal periods such as elections.

c) Both automated and human moderation must be available in multiple languages and be able to contextualise and algorithmically demote hate speech to ensure it does not become virulent prior to removal. Social media companies to carry out human rights impact assessments of its AI systems in place.

d) Collaborate with civil society, researchers and professional organisations on initiatives to counter hate, extremism, misinformation and disinformation. Transparency reports should include data on all content moderation and removal relating to hate speech, including those removed on the basis of government requests.

e) Consult researchers and CSOs in developing new tools and approaches for detecting and combating CIB.

#### **4. Media organisations**

a) Media must guard against propagating hate speech and refrain from giving such politics a forum or platform. Continue ethical and responsible journalistic practices as the standard bearer of facts and the watchdog of democracy.

b) Challenge and expose binary frameworks (us vs. them) that are divisive and discriminatory; reporting to focus on exposing and calling out hate speech and to provide solutions to stop such narratives; and have more news and programmes during elections that emphasises saying “No to Hate Speech”.

c) Media can also amplify and provide counter or alternative positions to combat the proliferation of messages of intolerance or expressions which may incite violence, hostility, or discrimination based on race, religion, royalty, gender, LGBTIQ and refugees and migrants. This should also include reporting on different groups or persons who are often the targets of hate speech and giving their members the opportunity to speak and to be heard in a way that promotes a better understanding of their perspectives and experiences.

#### **5. KOL and influencers**

a) KOL and influencers play a significant role in shaping public discourse and choice. It is imperative that KOL are cognisant of and draw a line between legitimate freedom of expression and hate speech.

b) KOL and influencers should also take on a proactive role in condemning harmful and hate speech by others.

#### **6. Social media users**

a) Be more discerning when engaging with online content. Stakeholders such as governments and social media platforms have big roles to play in building a healthy online information ecosystem. But individual users can still play their part. For example, if a user is unsure about any information on a social media platform, they should not share or forward such content. They can ignore or leave a UGC questioning the content’s authenticity or urge others to use fact-checking sites. Users can also search for credible news sources to influence a platform’s algorithm, which in turn determines what appears on their social media feed.

b) If they or someone they know is being threatened with harm, violence or is subjected to doxing, report it to the social media platform. Even if the target is a stranger, demand that content moderators take down any content that calls for violence against another

individual or a group of people. Hold the platforms accountable to their own community guidelines.

## 7. Civil society

a) Advocate for voluntary regulation in the industry, guided by an ethical standard and an independent, transparent, and self-regulated mechanism which is open to engagements with stakeholders and accountable to the public. In some countries, there is a move to make social media platforms liable for the harm caused by algorithm amplification.

b) Utilise the hate speech framework initiated through this project. There is a need to strengthen the monitoring framework adopted by this project, specifically the severity threshold and its implementation. With consistency, there will be greater clarity that will inform how the thresholds are defined and used as a baseline for future studies.

c) Assist in the development of AI tools for social/political monitoring. Having gone through the process of identifying what was needed for this type of study, the project team calls on CSOs and academicians to work with developers on AI-like tools to enable similar research in the future. These tools can be trained, enabling them to identify relevant issues and nuanced content. They can also be made more user-friendly and accessible for monitoring issues, not just for market research.

d) Study the impact of hate speech. This would require face-to-face interviews, as well as looking at the phenomenon over time to identify immediate and long-term effects, including the political, social and economic cost of hate speech. These qualitative studies can explore causes and consider intersectional factors such as class, race, gender, geographical locations and education level, painting a more complex picture and understanding of hate actors, amplifiers, and people's reactions and how they interpret hate speech.



## ANNEX I: LIST OF ACCOUNTS

### MEDIA

1. BERITA HARIAN ONLINE
2. SINAR HARIAN ONLINE
3. ASTRO AWANI ONLINE
4. SIN CHEW DAILY
5. UTUSAN MALAYSIA ONLINE
6. MALAYSIAKINI
7. ORIENTAL DAILY
8. THE STAR ONLINE
9. HARAKAH
10. NST ONLINE
11. BERNAMA.COM
12. BORNEO POST
13. BFM89.9
14. DAYAK DAILY
15. ROCKET
16. THE MERDEKA TIMES

### POLITICAL PARTIES

1. DEMOCRATIC ACTION PARTY (DAP)
2. PARTI ISLAM SE-MALAYSIA (PAS)
3. UNITED MALAYS NATIONAL ORGANISATION (UMNO)
4. PARTI PRIBUMI BERSATU MALAYSIA (BERSATU)
5. MALAYSIAN CHINESE ASSOCIATION (MCA)
6. PARTI AMANAH NEGARA (AMANAH)
7. PARTI Keadilan Rakyat (PKR)
8. PARTI WARISAN SABAH (WARISAN)
9. PARTI PEJUANG TANAH AIR (PEJUANG)
10. MALAYSIAN UNITED DEMOCRATIC ALLIANCE (MUDA)
11. GABUNGAN PARTI SARAWAK (GPS)
12. GABUNGAN RAKYAT SABAH (GRS)
13. BERJASA

## POLITICIANS

AMANAH – DATUK SERI HAJI MOHAMAD SABU

BERSATU – TAN SRI DATO' MUHYIDDIN @ MAHIADDIN YASIN

BERSATU – DATO' SERI MOHAMED AZMIN ALI

BERSATU – DATUK SERI RINA HARUN

BERSATU – DATUK SERI PANGLIMA HAJI MASIDI MANJUN

DAP - LIM GUAN ENG

DAP - HANNAH YEOH

DAP - TERESA KOK

DAP - ANTHONY LOKE

DAP - DR P. RAMASAMY

GPS – DATO' SRI NANCY SHUKRI

MCA – DATUK SERI WEE KA SIONG

MUDA - SYED SADDIQ SYED ABDUL RAHMAN

MUDA - AMIRA AISYA ABD AZIZ

MUDA - DR THANUSSHA FRANCIS XAVIER

MUDA – DR MATHEN NAIR

MUDA – LIM WEI JIET

PAS – TAN SRI DATO' SERI TUAN GURU HAJI ABDUL HADI AWANG

PAS – DATO SERI TAKIYUDDIN HAJI HASSAN

PAS – DATO' SERI MUHAMMAD SANUSI MD NOR

PBM – DATUK ZURaida KAMARUDDIN

PEJUANG – TUN DR MAHATHIR MOHAMAD

PKR – DATO' SERI ANWAR IBRAHIM

PKR - NURUL IZZAH ANWAR

PKR – MOHD RAFIZI RAMLI

PKR – DR MASZLEE MALIK

PKR – DATUK SERI SAIFUDDIN NASUTION ISMAIL

UMNO – DATO' SRI MOHAMMAD NAJIB TUN HAJI ABDUL RAZAK

UMNO – DATO' SERI DR AHMAD ZAHID HAMIDI

UMNO - KHAIRY JAMALUDDIN ABU BAKAR

UMNO – DATO' SRI AZALINA OTHMAN SAID

UMNO – DATUK SERI BUNG MOKHTAR RADIN

UMNO – DATUK SERI HJ AHMAD MASLAN

UMNO – DATUK SERI ABDUL AZEEZ ABDUL RAHIM

WARISAN – DATUK SERI PANGLIMA MOHD SHAFIE APDAL

UMNO – DATO' SRI ISMAIL SABRI YAAKOB

MIC – TAN SRI S. VIGNESWARAN

MIC – DATUK SERI M. SARAVANAN

UMNO – DATO' SERI UTAMA HAJI MOHAMAD HASAN

GPS – DATO WEE HONG SENG

GPS - LUKANISMAN AWANG SAUNI

GPS – DATO' SRI FADILLAH YUSOF

GPS – DATO GERALD RENTAP JABU

DAP – DR KELVIN YII

DAP - CHONG CHIENG JEN

DAP - VIOLET YONG

GPS - DR SIM KUI HIAN

DAP – GOBIND SINGH DEO

DAP - YOUNG SYEFURA OTHMAN

#### KEY OPINION LEADERS

1. HRH CROWN PRINCE OF JOHOR

2. MUHAMMAD ZAMRI VINOOTH KALIMUTHU

3. FIRDAUS WONG WAI HUNG

4. GERAKAN PENGUNDI SEDAR

5. FRIENDS OF IMMIGRATION

6. IKATAN MUSLIMIN MALAYSIA (ISMA) PRESIDENT AMINUDDIN YAHAYA

7. PERTUBUHAN-PERTUBUHAN PEMBELA ISLAM

8. AKADEMI DR HAR (DR HARLINA HALIZAH SIRAJ)

9. DR RAFIDAH HANIM MOKHTAR

10. CENTRE FOR HUMAN RIGHTS RESEARCH AND ADVOCACY (CENTHRA)

11. NURULHIDAYAH AHMAD ZAHID

12. NAMEWEE

13. NUMAN AFIFI

14. NUR FATHIAH SYAZWANA SHAHARUDDIN (CLEOPATRA)

15. SISTERS IN ISLAM

16. JUSTICE FOR SISTERS

## GOVERNMENT AGENCIES

1. POLIS DIRAJA MALAYSIA (PDRM)
2. JABATAN KEMAJUAN ISLAM MALAYSIA (JAKIM) OR DEPARTMENT OF ISLAMIC DEVELOPMENT MALAYSIA
3. IMMIGRATION DEPARTMENT OF MALAYSIA





CENTRE FOR  
**INDEPENDENT  
JOURNALISM**

Advocating Media Freedom and Access to Information