

Exploring the Social and Cultural Implications of Hate Speech on Facebook in Malaysia

New Kuwi Hoi*, Tay Shi Joh

Department of Journalism & Communication Studies, Faculty of Humanities and Social Sciences, Southern University College, 80300 Johor Bahru, Johor, Malaysia

*Corresponding author: khnew@sc.edu.my

Article history: Received: 10 April 2023 Received in revised form: 21 August 2023 Accepted: 11 September 2023 Published online: 31 December 2023

Abstract

The study provides insights into the prevalence of hate speech on Facebook among Malaysian users and highlights the need for greater efforts to counteract hate speech in Malaysia. Effective policies and interventions could be developed to address this issue based on the findings. This study examines hate speech trends among Facebook users in Malaysia and their attitudes towards regulating such speech. The study employs a four-part survey questionnaire: Part A covers demographics (gender, age, education); Part B examines social media engagement (platform choice, Facebook usage); Part C explores Facebook hate speech (locations, types, frequency); and Part D investigates responses (experiences, emotions, actions, perspectives). It was conducted with 225 respondents. The study found that race-related content is not the most commonly observed form of hate speech on Facebook in Malaysia; instead, users commonly encounter other types of hate speech. Additionally, users' level of awareness regarding hate speech is influenced by their level of activity on the platform. Notably, the duration of time spent on Facebook is not correlated with the frequency of exposure to hate speech. Experiences of being victimized by hate speech do not necessarily lead to support for regulating hate speech on Facebook. These findings provide valuable insights into the prevalence of hate speech on Facebook among Malaysian users and underscore the need for greater efforts to counteract hate speech in Malaysia. The study suggests the importance of developing effective policies and interventions to address this issue based on the findings.

Keyword: Hate speech, Facebook, Malaysia, Regulation, Prevalence

Abstrak

Kajian ini mengutarakan tentang kelaziman isu ucapan kebencian dalam kalangan pengguna Facebook di Malaysia dan menekankan tentang perlunya suatu wawasan dalam menangani isu ucapan kebencian di Malaysia. Kajian ini mengkaji trend ucapan kebencian dalam pengguna Facebook di Malaysia dan sikap mereka terhadap mengawal selingan sedemikian. Kajian berdasarkan empat bahagian soal selidik: Bahagian A merangkumi demografi (jantina, umur, pendidikan); Bahagian B mengkaji penglibatan media sosial (pilihan platform, penggunaan Facebook); Bahagian C menerokai ucapan kebencian dalam Facebook (lokasi, jenis, kekerapan); dan Bahagian D menyelidiki maklum balas (pengalaman, emosi, tindakan, pandangan). Seramai 225 responden ditemui, hasil kajian mendapati bahawa kandungan berkaitan perkauman bukanlah bentuk ucapan kebencian yang paling lazim terdapat dalam Facebook di Malaysia; sebaliknya, pengguna kerap berdepan dengan jenis ucapan kebencian bersifat peribadi. Selain itu, tahap kesedaran pengguna mengenai ucapan kebencian dipengaruhi oleh tahap aktiviti mereka dalam platform tersebut. Perlu dinyatakan bahawa tempoh masa yang dihabiskan di Facebook tidak menunjukkan kolerasi dengan kekerapan pendedahan kepada ucapan kebencian. Pengalaman menjadi mangsa ucapan kebencian tidak semestinya membawa kepada sokongan terhadap mengawal ucapan kebencian dalam Facebook. Dapatan ini memberikan wawasan berharga mengenai prevalensi ucapan kebencian dalam Facebook di kalangan pengguna Malaysia dan menekankan keperluan wawasan untuk menangani isu ucapan kebencian di Malaysia. Berdasarkan penemuan, kajian mencadangkan kepentingan membina dasar dan intervensi yang berkesan untuk menangani isu ini.

Kata Kunci: Ucapan Kebencian, Facebook, Malaysia, Kawalan, Kelaziman

© 2024 Penerbit UTM Press. All rights reserved

1.0 INTRODUCTION

Along with the advancement of technology, the Internet has become popular and contributed too many innovations, including social media. People rely on social media for information, entertainment, and communication. According to Statista (2022), Malaysia had ninety-one percent of social media users, from the total population. And according to the latest ranking of the top fifteen social media platforms and applications, Facebook is ranked at number one. It has the largest number of active users, with around two billion seven hundred forty million users worldwide (Karl, 2023). According to Karl (2023), Facebook is a free social networking website that allows registered users to build profiles, post photos and videos, send messages, and stay in touch with friends, family, and co-workers. The most recent data in December 2022 shows Malaysia has around thirty-one million Facebook users, accounting for around ninety percent of the total population (Napoleon Cat, 2022).

The contradiction between the restriction of hate speech and free speech has always existed. People are encouraged to voice their opinions, especially in those countries that emphasize democracy, such as Malaysia. It is considered a public right where people are free to express their ideas. However, Guiora and Park (2017) stated that people have gradually neglected their due social responsibilities by

relying on the so-called right to freedom of speech. Hate speech that is offensive and insults others has gradually risen and received attention. The damage caused by overly free or even extreme speech is invisible but not less than the actual damage, which poses a threat to society. In the books written by Cohen-Almagor, "Confronting the Internet's Dark Side: Moral and Social Responsibility on the Free Highway (2015)," he advocates a balance between freedom of speech and social responsibility. He mentioned that hate speech has a connection with hate crime by citing examples such as inflammatory Tweet from a Sam Houston State University student. Seeing the harm that these hate speeches cause to society, the government has set restrictions on speech to reduce the possible harm caused by speech.

In Malaysia, the people have freedom of speech under the restrictions of the law under the Federal Constitution of Malaysia. According to Article 10 of the Federal Constitution of Malaysia, it mentions the regulations on freedom of speech, assembly, and association. Even though Malaysians have the right to freedom of speech, there are some restrictions stated in Article 10 where legal action will be taken once anyone breaks the rules (Law, 2013). Under Article 10, the latest amendment was made to the Sedition Act of 1948. It states that the promotion of hatred between religious groups is not permitted. It is necessary to restrict hate speech, as Malaysia is a multi-racial country. The impact of breaking the rules, on the other hand, is not being paid much attention to and is not even noticing by society. There are no specific acts towards hate speech under the Federal Constitution of Malaysia, unlike other countries such as Germany and France that have specific laws to address hate speech (Nor & Ab Razak, 2017). In accordance with the stipulations set forth in Penal Code 504 and 505, the act of deliberately causing insult with the intention to incite a breach of the peace is subject to legal consequences, encompassing imprisonment for a duration of up to two years, imposition of a fine, or a combination of both punitive measures (Act 574, Penal Code As at 1 February 2018, pg 263). Furthermore, under the provisions of the Communications and Multimedia Act 1998, individuals found to have committed an offense falling within the purview of sections 233 (1) (a) & (b) and 233 (2) (a) & (b) are subject to conviction, rendering them liable to a fine not exceeding fifty thousand ringgit, a potential imprisonment term not surpassing one year, or a concurrent application of both penalties. Additionally, such offenders will also be susceptible to an incremental fine of one thousand ringgit for each day during which the offense persists subsequent to the pronouncement of conviction (Communications and Multimedia Act 1998, Act 588). Therefore, the necessity of regulating hate speech with a specific law in Malaysia still requires more opinions and confirmation according to the remaining variables.

This study aims to investigate how often Malaysians see hate speech on Facebook and draw attention to hate speech in Malaysia. Race, religion, and royal institutions are sensitive issues that must be dealt with care in a multi-ethnic and multicultural society like Malaysia.

■ 2.0 LITERATURE REVIEW

According to Brown, the term "hate speech" is ambiguous, and it implies a range of connotations. He defines "hate speech" as more than just a controversial phrase; it's also "systematically ambiguous. (Brown, 2017, as cited in Anderson & Barnes, 2022). That is, it encompasses multiple different interpretations." Since the phrase is a fundamentally contested term, searching for a single or universal definition is futile (Lua (2022), Professor Cohen-Almagor sees hate speech weakens or makes the effort of preserving this public good considerably more difficult than it would otherwise be. It accomplishes this not just by implying prejudice and brutality, but also by reawakening live nightmares of what this society—or other societies—were like in the past (Waldron, 2012, as cited in Guiora & Park, 2017). Hate speech definitions based on intrinsic features often relate to those that stress the sort of speech delivered. What is at stake is the bulk of society's usage of speech that is commonly recognized to incite offence or insult. Slurs and other explicitly disparaging terms are classic instances of this point of view. In general, the speech identified on this account is disparaging, discriminatory, or vilifying in nature.

The openness of social media has also allowed people to achieve the goal of free speech, however, this is detrimental to society. Free speech is the freedom to freely express one's views and words, while hate speech is a remark that abuses this right to injure others or incite violence (Buang, 2019, as cited in Mia et al. 2021). Guiora and Park (2017) mentioned that online social media is a platform where most disputes and misunderstandings arise from freedom of speech. The network is unlimited and has great space and influence. The rights provided to the people by many countries that emphasize democracy include freedom of speech. Even if speech is divided into legally protected and unprotected, freedom of speech reflects the health and responsibility of society, but extreme speech is challenging society.

According to the findings of Guo and Johnson (2020), users who read hate speech related to racism, sexism, and Lesbian, Gay, Bisexual, and Transgender (LGBT) on Facebook perceive the effects on the general public to be greater than their own. While paternalism does not affect the effect perceiving of the users. Also, the result shows that female users are more likely to report sexism hate speech as they perceive the effects on general public. They are more willing to support censorship of Facebook. Compared to sexism, hate speech, racism and LGBT hate speech does not affect the user's behavior. Lastly, the support for freedom of speech do not relate to the perceiving of effects on general public and own self (Guo & Johnson, 2020).

A tragedy happened in 2017, 20-year-old Teh Wen Chun, jumped from a 17th-floor flat in Tanjung Bungah, Georgetown, Penang, and leapt to his death. He became depressed after reading an article that sullied his internet reputation. His father told The Star Online that he was aware of his son being abused on the internet. Wen Chun's behaviour changed when several of his college friends insulted and called him names on Facebook. However, Wen Chun did not express any feelings or complain to his father (Ariffin et al., 2021). "As a result of continued abuse, they may suffer from depression which could lead to a host of other physical and mental problems." (Nor & Ab Razak, 2017). The consequences of hate speech are invisible, and it affects people psychologically with repercussions that gradually affect social development. According to the University Putra Malaysia government and civilizational studies department senior lecturer Murni Wan Mohd Nor, she continued by stating hate speech may be harmful to the healthy connections of society's members of all races and religions (Nor & Ab Razak, 2017).

In fact, some scholars of hate speech have discovered the potential consequences of hate speech regarding various human values. Ruwandika (2018) mentioned that hate speech is highly correlated with freedom of expression, individual, group and minority rights and concepts of dignity, liberty, and equality. With this expansion, platforms like WhatsApp, Twitter, and Facebook have seen an increase in

online hate speech, one sort of which is prejudice based on race, ethnicity, and heritage. And in other instances, this discourse has led directly to negative effects in the real world, such as crime and violence. Racist hate remarks began to dominate social media after the "tabligh cluster" was detected during a religious gathering in Sri Petaling, Kuala Lumpur in February, accounting for nearly half of the total Covid-19 cases in Malaysia. Internet users then blamed the recent increase in the number of new Covid-19 cases on refugees, immigrants, and foreign employees (Hassandarvish, 2020).

Posting false information about the victim, making sexual remarks, threatening them, or using offensive terms like hate speech are all examples of cyberbullying. However, the line between offensive language, hate speech, and cyberbullying is frequently unclear. Due to new tools and methods of assault used online, cyberbullying is developing, making it difficult to define precisely. When victims are verbally abused through remarks made online, in addition to being harassed, cyberbullying becomes more difficult (Xu, 2021). Furthermore, neither of these expressions contribute anything positive to society. Both include the risk of causing pain, violence, and death, whether to oneself or to others. In both cyberbullying and hate speech, the victims are meant to suffer as a result of the communication. It is true that a lot of cases of online bullying and hate speech have sparked violent behaviour and hate crimes (Cohen-Almagor, 2022). There are four hypotheses that have been proposed in this study, there are:

- H1 Races content is the most seen hate speech on Facebook.
- H2 There is a relationship between activeness on Facebook and the awareness of hate speech on Facebook.
- H3 There is a relationship between hours spend on Facebook and the frequency of seeing hate speech on Facebook.
- H4 There is a relationship between experiences of hate speech attack and support of Facebook's regulation.

According to the Council of the European Union, "hate speech" encompasses a wide range of expressions that serve to propagate, incite, endorse, or rationalize racial animosity, xenophobia, anti-Semitism, and other forms of intolerance. This includes the promotion of extreme nationalism, ethnocentrism, discrimination, and hostility towards minority groups, migrants, and individuals with immigrant backgrounds (Justitia, 2020). Furthermore, the Council of Europe extends its purview to encompass hate speech and discrimination across various domains, such as education, sports, Roma rights, gender equality, sexual orientation, gender identity, media, and internet regulation (Council of Europe, 2022). Within the context of this study, we define hate speech as any expression on Facebook, whether in the form of posts or comments, that is crafted with the intent to foster hatred, launch attacks, offer criticism, or engage in insults directed at individuals with regard to general attributes, physical appearance, race, political affiliations, or involvement in sports. These dimensions represent common topics of discussion and engagement for people on the platform. Thus, this research centres on examining the prevalence of such forms of hate speech within the Malaysian Facebook community.

■ 3.0 METHODOLOGY

This study employed a quantitative survey design and utilized online platforms to distribute questionnaires to Facebook users. Purposive sampling was adopted to target a specific issue related to the trend of hate speech on Facebook among Malaysians. The study investigated several variables, including the time spent on Facebook, the frequency of encountering hate speech, and the most commonly encountered types of hate speech. The population of interest for this research was Malaysian Facebook users. Purposive sampling seeks to specifically target Malaysian netizens who actively use Facebook, spanning an age range from 18 to 57 years old as its defining criteria. The sample consisted of 225 respondents. Initially, the researcher provides a concise differentiation of the study's population and sample. The study specifically focuses on the populace of Facebook users within Malaysia, and the researcher's target pertains exclusively to this cohort. Subsequently, the delineated population is methodically stratified into smaller clusters predicated on age classifications, culminating in the identification of four distinct subgroups from the initial population selection.

This study has demonstrated a strong commitment to addressing ethical considerations within the context of the sensitive subject matter of hate speech. As a result, no personal identifiers such as names, races, or religions were solicited from the respondents, a deliberate approach aimed at mitigating any potential discomfort or unease among the respondents. Drawing from the guidelines set forth by the Australian Bureau of Statistics, which advocates for age groupings of either five or ten years for optimal utility across diverse applications (Age Standard, 2014), the present study adopts the age range categorization of "10s." This discernment leads to the establishment of four distinct groups encompassing individuals aged 18 to 27 years, 28 to 37 years, 38 to 47 years, and 48 to 57 years. According to Bennekom (2018), when dealing with a substantial population, the inclusion of 200 respondents can yield a remarkably accurate survey outcome, aligning with the assumptions and constraints inherent in survey-based investigations. This empirical insight underscores the efficacy of such an approach within the study's framework. After collecting responses from the participants, the researcher thoroughly reviewed and checked the data for completeness and accuracy. Any missing values were removed to ensure the credibility of the results. The data was then extracted into an excel file and entered into SPSS software for analysis. The responses to Likert scale questions were coded based on the degree of agreement, ranging from "very rarely" coded as 1, to "very often" coded as 6. The researcher followed an ascending coding method, from "agree" to "disagree" and from "less" to "more," to ensure consistency in the data analysis process. Overall, this careful data management process aimed to guarantee the accuracy and reliability of the study's findings. The questionnaire, carefully crafted to closely align with our research objectives and subjected to a rigorous pilot test that engaged 30 randomly selected individuals not belonging to the primary research population, as outlined by Chua (2020). Furthermore, the questions within Part C, focusing on Hate Speech, incorporating Likert scale responses, underwent a comprehensive reliability analysis utilizing Cronbach's Alpha reliability statistics. The obtained result, with a Cronbach's Alpha of .777, signifies a commendable level of reliability and acceptability, as detailed in the table 1 and table 2 below.

Table 1 Reliability Test for questionnaire

Reliability Statistics			
Cronbach's Alpha	Cronbach's Alpha Standardized	Based on	N of Items
.777	.788		4

Table 2 Reliability Test Details for Part C of the Questionnaire

Item Statistics			
	Mean	Std. Deviation	N
I feel angry when I see hate speech on myself	3.85	.902	225
I feel angry when I see hate speech on family members	4.45	.725	225
I feel angry when I see hate speech on friends	4.25	.798	225
I feel angry when I see hate speech on others	3.53	.973	225

4.0 RESULTS

4.1 Descriptive Analysis

Table 3 presents a comprehensive overview of our analysis, unveiling insightful demographic details of the study participants. Among the social media users surveyed, a significant 88.0% fall within the 18-27 age bracket, followed by 7.1% in the 28-37 age group, 2.7% in the 38-47 cohort, and 2.2% in the 48-57 category. Furthermore, our analysis underscores a gender distribution, with 74.7% identifying as females and 25.3% as males. In terms of educational attainment, the majority of respondents, accounting for 55.1%, hold a degree certificate. Additionally, 12.9% possess a diploma, 13.8% have completed Foundation or Sijil Tinggi Persekolahan (STPM), 16.0% possess a middle high school qualification (SPM/UEC), and 2.2% have achieved a master's degree.

Table 3 Demographic Analysis of Respondents

		Frequency	Percent
Gender	Female	168	74.7
	Male	57	25.3
	Total	225	100.0
Age Groups	18 to 27	198	88.0
	28 to 37	16	7.1
	38 to 47	6	2.7
	48 to 57	5	2.2
	Total	225	100.0
Education Levels	Degree level	124	55.1
	Diploma level	29	12.9
	Foundation/STPM level	31	13.8
	Master	5	2.2
	SPM level	36	16.0
	Total	225	100.0
The Purpose of Using Social Media	All	2	0.9
	Business	6	2.7
	Communication	63	28.0
	Entertainment	129	57.3
	Information	23	10.2
	Information, Com & Entertainment	2	0.9
	Total	225	100.0

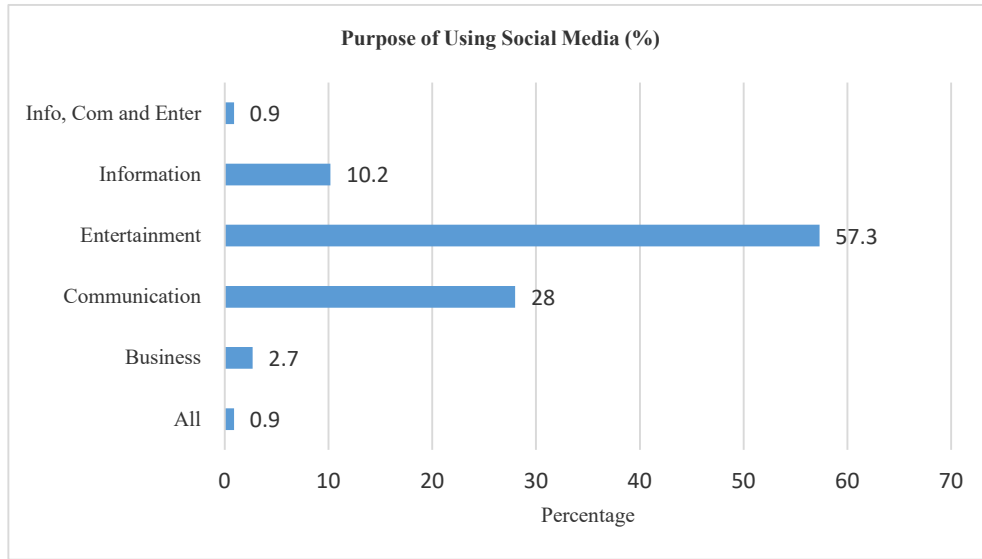


Figure 1 Purpose of Using Social Media (%)

Figure 1 shows there are 57.3% of respondents who use Facebook for entertainment purposes. For the first choice of social media, there are 48.9% of respondents choose Instagram (as shown in figure 2). While the total of 37.3% respondents choose Instagram as their second choice of social media in Figure 3. Figure 4 shows 56.9% of respondents spend 0-2 hours on Facebook per day.

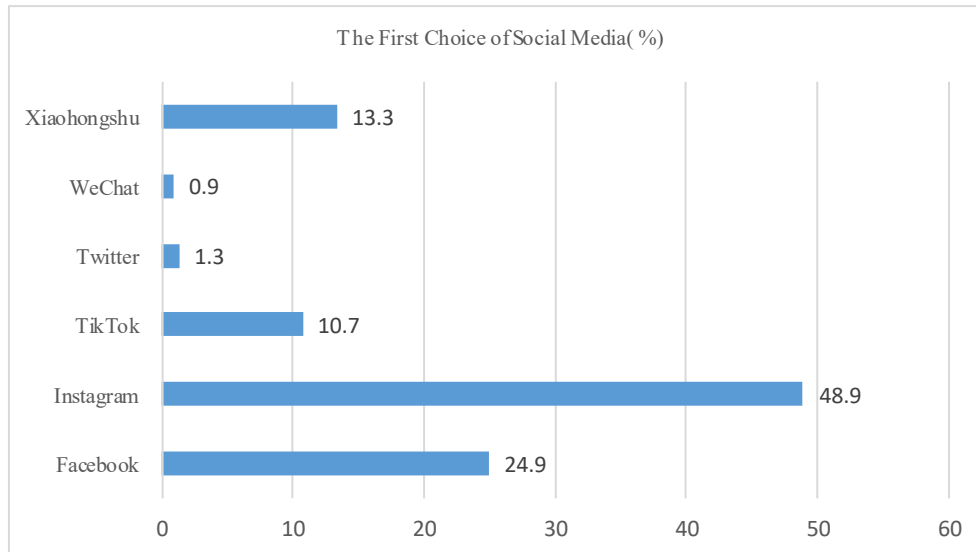


Figure 2 The First Choice of Social Media (%)

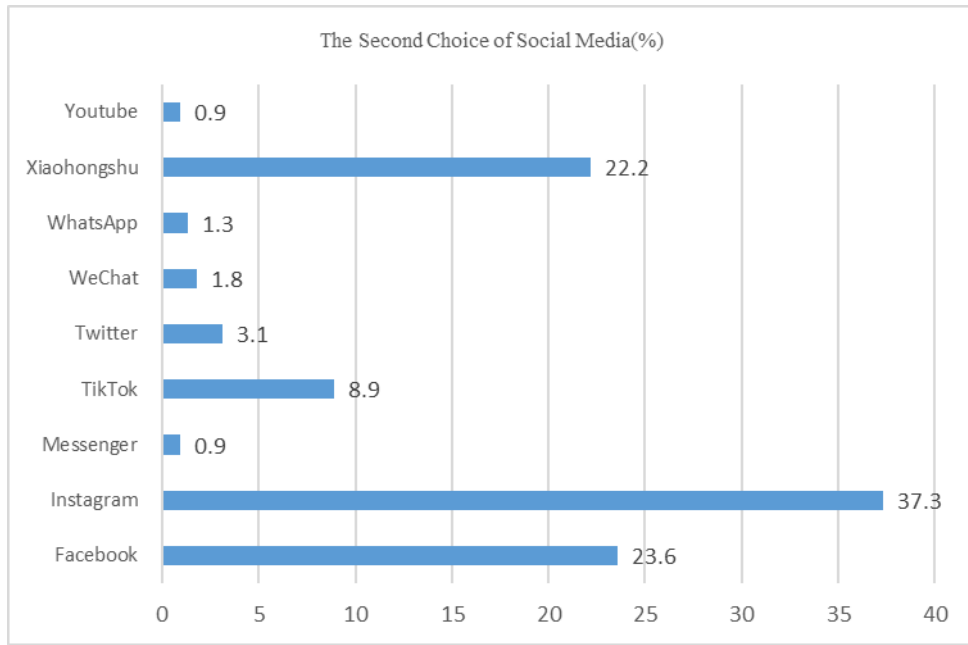


Figure 3 The Second Choice of Social Media (%)

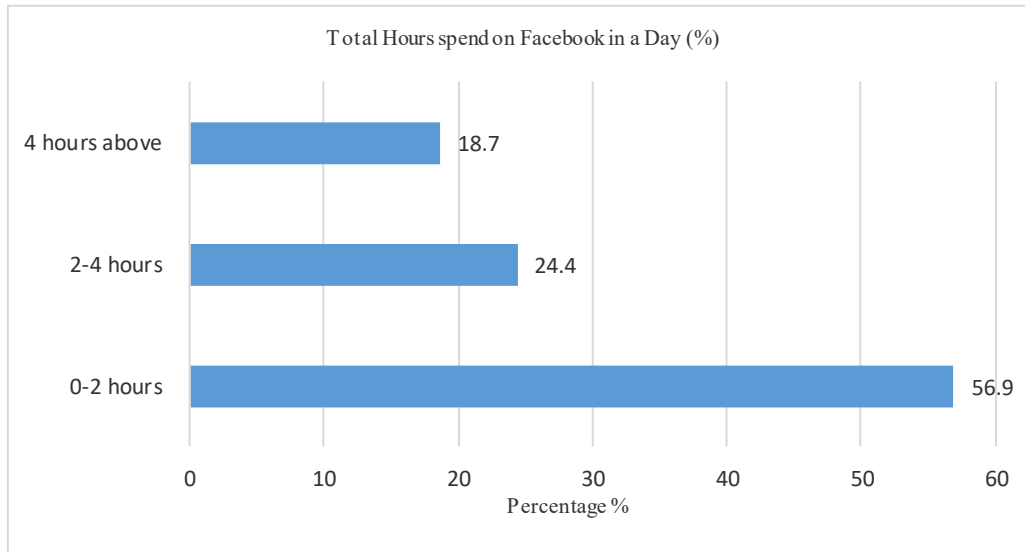


Figure 4 Total Hours Spend on Facebook in a Day (%)

Table 4 Awareness on Hate Speech

		Frequency	Percent
Awareness on Hate Speech	Strongly disagree	54	24.0
	Disagree	95	42.2
	Neutral	58	25.8
	Agree	11	4.9
	Strongly agree	7	3.1
	Total	225	100.0
Where I See Hate Speech the Most	Facebook comments	209	92.9
	Facebook posts	16	7.1
	Total	225	100.0
Frequency of Seeing Hate Speech	Strongly disagree	7	3.1
	Disagree	15	6.7

	Somewhat disagree	39	17.3
	Somewhat agree	78	34.7
	Agree	61	27.1
	Strongly agree	25	11.1
	Total	225	100.0
The Most Seen Hate Speech Content	All	2	.9
	General hate	63	28.0
	Personal attack	106	47.1
	Political	26	11.6
	Races	20	8.9
	Sports	8	3.6
	Total	225	100.0

Table 4 shows among a total 42.2% of respondents, disagree that they know about hate speech. 92.9% of respondents saw hate speech in Facebook comments. 34.7% of respondents somewhat agree that they often see hate speech on Facebook. Among the majority of respondents, 47.1% see personal attack hate speech commonly.

Table 5 Responses on Seeing Hate Speech about oneself, family, friends and others

		Frequency	Percent
I Feel Angry when I See Hate Speech about Myself	Strongly disagree	4	1.8
	Disagree	7	3.1
	Neutral	65	28.9
	Agree	91	40.4
	Strongly agree	58	25.8
	Total	225	100.0
I Feel Angry when I See Hate Speech about Family Members	Strongly disagree	1	.4
	Disagree	1	.4
	Neutral	22	9.8
	Agree	72	32.0
	Strongly agree	129	57.3
	Total	225	100.0
I Feel Angry when I See Hate Speech about Friends	Strongly disagree	1	.4
	Disagree	6	2.7
	Neutral	26	11.6
	Agree	94	41.8
	Strongly agree	98	43.6
	Total	225	100.0
I Feel Angry when I See Hate Speech about Others	Strongly disagree	10	4.4
	Disagree	13	5.8
	Neutral	85	37.8
	Agree	82	36.4
	Strongly agree	35	15.6
	Total	225	100.0

Table 5 shows 40.4% of respondents agree that they feel angry when seeing hate speech on their own selves. 57.3% of respondents strongly agree that they feel angry when seeing hate speech about their family members. 43.6% of respondents strongly agree that they feel angry when seeing hate speech about their friends.

Table 6 Experience and Action Taken on Hate Speech

		Frequency	Percent
I Had Experienced Hate Speech Attack on Facebook	Strongly disagree	139	61.8
	Disagree	50	22.2
	Neutral	20	8.9
	Agree	14	6.2
	Strongly agree	2	.9
	Total	225	100.0
Action on Hate Speech	Block	47	20.9
	Ignore	120	53.3
	Report to Facebook	58	25.8
	Total	225	100.0
It Is Necessary to Regulate Hate Speech on Facebook	Strongly agree	78	34.7
	Agree	39	17.3
	Somewhat agree	39	17.3
	Somewhat disagree	22	9.8
	Disagree	23	10.2
	Strongly disagree	24	10.7
	Total	225	100.0
It Is Necessary to Raise Awareness on Hate Speech Issues in Malaysia	Strongly agree	117	52.0
	Agree	35	15.6
	Somewhat agree	20	8.9
	Somewhat disagree	25	11.1
	Disagree	27	12.0
	Strongly disagree	1	.4
	Total	225	100.0

Table 6 shows 61.8% of respondents strongly disagree that they had experienced hate speech attacks on Facebook. 53.5% of respondents ignore hate speech on Facebook. 34.7% of respondents strongly agree to regulate hate speech on Facebook. 52.0% of respondents strongly agree to raise awareness on hate speech issues in Malaysia.

Hypothesis Testing

H1: Racial Content Is the Most Commonly Seen Hate Speech on Facebook.

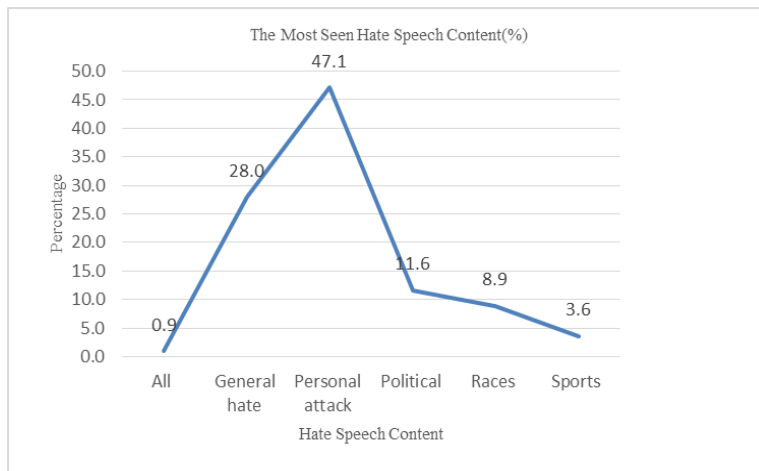


Figure 5 Line Chart of Most Seen Hate Speech Content

The result as shows in Figure 5 indicates the content of hate speech, personal attack 47.1%, general hate 28.0%, political 11.6%, races 8.9%, sports 3.6%, and all content 0.9%. H1 is rejected where the racial content is not the most seen hate speech on Facebook, but personal attacks content was commonly seen instead.

Table 7 Simple Linear Regression Analysis on Most Seen Hate Speech Content

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.649	1	3.649	3.349	.069 ^b
	Residual	243.000	223	1.090		
	Total	246.649	224			
a. Dependent Variable: The most seen hate speech content						
b. Predictors: (Constant), Where you see hate speech the most						

The researcher further analyses the relationship between where the users see hate speech the most and the most seen hate speech content by using simple linear regression. (Table 7) There is no significant change in the most commonly seen hate speech content due to where the users see hate speech the most, because of the significant value is (sig=.069), which is more than the acceptable value (p value<.05).

H₂: There Is A Relationship Between Activeness on Facebook and the Awareness of Hate Speech on Facebook.

Table 8 Pearson Chi Square Analysis on Hate Speech

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	43.089 ^a	20	.002
Likelihood Ratio	48.924	20	.000
Linear-by-Linear Association	4.765	1	.029
N of Valid Cases	225		
a. 15 cells (50.0%) have expected count less than 5. The minimum expected count is .34.			

A Pearson Chi-Square test was performed to test the association between activeness on Facebook and awareness of hate speech on Facebook.(Table 8) Pearson Chi-Square test revealed that the association between two variables is statistically significant, as Asymptotic Significance (2-sided), p< .05. An association between activeness on Facebook and awareness of hate speech on Facebook was observed, $\chi^2(20) = 43.09$, p = .002. Hence, the null hypothesis is rejected while H₂ is accepted.

H₃: The More Hours Spent on Facebook, The Higher The Frequency of Seeing Hate Speech on Facebook.

Table 9 Simple Regression Analysis on Frequency of Seeing Hate Speech and hours spent on Facebook

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.210	1	4.210	2.963	.087 ^b
	Residual	316.830	223	1.421		
	Total	321.040	224			
a. Dependent Variable: Frequency of seeing hate speech						
b. Predictors: (Constant), The hours I spend on Facebook per day						

A simple linear regression was performed to test the effect of hours spent on Facebook, on the frequency of seeing hate speech on Facebook.(Table 9). There is no significant change in the frequency of seeing hate speech on Facebook due to the hours spent on Facebook, because the significant value is ($p=.087$), which is more than the acceptable value ($p\text{-value}<.05$). Hence, the null hypothesis is accepted while H_3 is rejected.

H4: There Is A Relationship Between Experiences on Hate Speech Attack and Support on Facebook Regulation.

Table 10 One-way ANOVA Analysis on Support of Facebook Regulation

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.653	1	1.653	.554	.458 ^b
	Residual	665.903	223	2.986		
	Total	667.556	224			
a. Dependent Variable: It is necessary to regulate hate speech on Facebook						
b. Predictors: (Constant), I have experience hate speech attack on Facebook						

A simple linear regression was performed to test the relationship between experiences on hate speech attack and support on Facebook regulation. (Table 10).There is no significant relationship because of the significant value ($p=.458$), which is more than the acceptable value ($p\text{-value}<0.05$). Hence, the null hypothesis is accepted while H_4 is rejected.

5.0 DISCUSSION AND RECOMMENDATION

H1: Racial Content Is The Most Seen Hate Speech on Facebook.

The study's findings have revealed an unexpected trend, contrary to our initial hypothesis (H1). It turns out that racial content is not the predominant form of hate speech on Facebook, which defies our expectations ($sig. =.069$; $p<.05$). Several factors may explain this disparity between our hypothesis and the actual results. These factors encompass potential limitations within our dataset, the ever-evolving content policies on Facebook, shifts in user behaviour, and other contextual influences that could have influenced our findings. As highlighted by Farkas et al. (2018, cited in Fernández and Farkas, 2021) and Hangartner et al. (2021), the policies and content moderation processes implemented by social media platforms, in conjunction with government regulations, exert a significant influence on the landscape of hate speech within these platforms.

In addition to these factors, our research encountered limitations concerning sample size, gender representation, and the choice of data collection platform, all of which might have contributed to the observed inconsistencies. These discrepancies underscore the intricate nature of defining 'hate speech,' suggesting a pressing need for further exploration. In line with Fernández and Farkas' findings (2021), which also point out disparities in the interpretation of concepts like racism and hate speech, it's evident that there's been limited scholarly exchange between the humanities and data science in this realm. Consequently, it's imperative that future research delves deeper into the factors that influence hate speech exposure on Facebook. While our hypothesis regarding racial content was not validated, the study's findings do indicate that personal attacks constitute the most prevalent form of hate speech on the platform. Despite this unanticipated result, our research offers valuable insights into the dynamics of hate speech on Facebook. It not only lays the groundwork for future investigations but also underscores the vital importance of continued research in this crucial area.

The results unmistakably demonstrate that personal attacks are the most commonly encountered form of hate speech on Facebook, rather than racial content. In a multi-ethnic country like Malaysia, sensitive issues such as race, royalty, and religion are highly classified. However, the discussion of these sensitive topics appears to be on the rise in recent years, likely fueled by technological advancements that afford people greater freedom and courage to broach these subjects. Between July 2020 and July 2022, the Malaysian Communications and Multimedia Commission (MCMC) recorded a total of 1,782 complaints related to hate speech associated with race, religion, and royal institutions (3R) (Malaysian Communications and Multimedia Commission (MCMC), 2022). To maintain social harmony, Malaysia must proactively combat hate speech, involving various ministries, agencies, and non-governmental organizations (NGOs) in these initiatives. Simultaneously, the existing legislative framework must be updated to ensure that Malaysia takes a comprehensive approach to addressing hate speech, including preventing it from escalating into incitement to discrimination, hatred, and violence, all of which are prohibited under international law (United Nations, 2019).

Nonetheless, this development is worrisome, as hate speech is the breeding ground for extremism and the first step toward radicalization, which can lead to violence. In a multi-ethnic and multicultural society like Malaysia, issues related to race, religion, and royalty (3R) require careful handling (Ahmad, 2020). However, our findings reveal that personal attack content is the most prevalent form of hate speech among Facebook users. This suggests that Malaysia, as a diverse country, might have a stereotype that racial content is the most significant issue that warrants attention. However, while racial hate speech can affect a group of people, personal attacks in hate speech can target individuals. Additionally, our results show that there is no significant relationship between where users encounter hate

speech and the specific type of hate speech they encounter. Facebook's openness does not seem to limit the exposure to hate speech. Based on our results, it appears that hate speech centered on personal attacks has been on the rise and is no longer limited to issues related to race, religion, and royal institutions (3R) that the government has primarily focused on.

H2: There Is A Relationship Between Activeness on Facebook and The Awareness of Hate Speech on Facebook.

A Pearson Chi-square analysis was performed to investigate the relationship between Facebook activity and awareness of hate speech. The results unveiled a statistically significant correlation between these two variables, ($\text{sig.} = .002, p < .05$). It's crucial to note that while statistical significance was achieved, this alone doesn't necessarily indicate a robust or meaningful association. These findings stress the importance of monitoring individuals' Facebook usage patterns. Specifically, it's imperative for individuals, organizations, and society as a whole to recognize that increased Facebook usage is linked to heightened awareness of hate speech incidents. Consequently, we should contemplate actions to regulate Facebook use, especially among young children and students.

In January 2022, Malaysia boasted a staggering ninety percent of Facebook users within its total population (NapoleonCat, 2022). Notably, WhatsApp, Facebook, and Instagram ranked among the top three most popular social media platforms in Malaysia (Rosley, 2022). The convergence of sensitive issues on social media platforms represents a significant threat to the nation. It's vital to underscore that when the source of hatred stems from an event outside the country, addressing the root cause becomes paramount. Therefore, there is a pressing need for greater efforts to nurture critical thinking among Malaysians, especially our youth. This entails building mental firewalls and enhancing digital literacy skills (Ahmad, 2020). Defining what it means to be "active" on social media isn't uniform; platforms themselves set the criteria. For instance, Facebook defines an active user as someone who has logged in and taken some form of action, like sharing or publishing, within the past 30 days (Visibility, 2016). Consequently, when users are actively engaged on Facebook, which entails viewing, sharing, and publishing content, they become more attuned to hate speech. This active engagement exposes them to a wider range of Facebook features, increasing their chances of encountering hate speech incidents.

H3: There Is A Relationship Between Hours Spend on Facebook and The Frequency of Seeing Hate Speech on Facebook.

In addition to our previous analysis, we employed a simple linear regression to determine whether the time spent on Facebook could predict the frequency of encountering hate speech. Initially, we hypothesized that increased time spent on social media, driven by the pursuit of information and entertainment, would lead to higher exposure to hate speech. It was anticipated that individuals who spent more time on social media would encounter hate speech more frequently. Research by Sert and Başkale (2022) supported this notion, revealing that prolonged social media use increased addiction scores among students, especially during the pandemic. In essence, more time spent on social media resulted in greater addiction. Furthermore, a study by Harriman et al. (2020) demonstrated that as children spent more time online, their likelihood of encountering hate speech increased.

However, this study contradicts these expectations, as the hours spent on Facebook showed no significant relationship with the frequency of encountering hate speech ($\text{sig.} = .087, p < .05$). Consequently, we accepted the null hypothesis and rejected H3. It's worth noting that several variables may influence these results, such as Facebook's content distribution, the type of content users typically engage with, and individual sensitivity to hate speech detection. These findings align with research by Kormelink and Meijer (2019), which indicated that the time spent on a platform does not necessarily reflect the quality of attention or the level of interest and engagement. In some cases, more time spent may indicate a lack of interest or involvement. For future research, it's advisable to consider introducing control variables to enhance accuracy. Factors like Facebook's content distribution, the type of content consumed, and individual sensitivity to hate speech could be critical in understanding hate speech exposure. Notably, sensitivity to detecting hate speech varies among users; some may not refer to Facebook comments, while most participants in our study reported encountering hate speech more frequently within these comments. These findings are consistent with previous research by Chetty and Alathur (2018), who defined hate speech as offensive communication conveying hate-based ideologies through stereotypes and targeting protected characteristics such as gender, religion, race, and disability. To effectively combat hate speech, they proposed the implementation of national and international legal frameworks and policies aimed at countering both hate speech and terrorism.

H4: There Is A Relationship Between Experiences on Hate Speech Attack and Support on Facebook Regulation.

The examination of the relationship between experiences of hate speech attacks and support for Facebook's regulation was conducted through simple linear regression analysis using SPSS software. Our hypotheses were framed as follows: the null hypothesis posits that there is no connection between experiences of hate speech attacks and support for Facebook's regulation, while H4 proposes that such a relationship exists. A similar hypothesis was explored by Guo and Johnson (2020), who discovered that support for regulation varied depending on the type of hate speech. According to the theory of third-person effects, individuals tend to perceive a greater impact on others, which may influence their support for regulation. They are predicted to be more inclined to support Facebook's regulation of hate speech.

In this study, a prediction that individuals who have never experienced a hate speech attack might perceive a greater impact on others, leading to support for Facebook's regulation. However, the results from the simple linear regression analysis indicate that experiences of hate speech attacks and support for Facebook's regulation do not have a significant relationship ($\text{sig.} = .458, p < .05$). Consequently, H4 is rejected, and the null hypothesis is accepted. Contrary to expectations, participants who have never experienced hate speech tend to be more willing to support Facebook's regulation. They are more likely to endorse hate speech regulation if they believe it has a greater impact on others. This finding contradicts the findings of Guo and Johnson (2020), who explored how perceived effects of Facebook hate speech influenced participants' subsequent censorial behaviour, suggesting that women were more likely to support censorship when they

perceived sexist speech as having a greater impact on other users. Nonetheless, it is argued that the type of hate speech plays a pivotal role in influencing support for Facebook's regulation. The nature of hate speech, whether it pertains to gender discrimination, race, politics, or other factors, is contingent on individual concerns. For instance, women may be more inclined to report hate speech and gender discrimination that targets them. Additionally, the extent of dissatisfaction with the content may affect their support for regulation. Some users might perceive the hate speech they encounter as relatively benign and thus deem regulation unnecessary. Furthermore, the concept of self-construal can also impact the third-person effect. Some individuals view themselves as part of a collective rather than as unique, independent entities, which is known as a collectivist self-construal. Interestingly, adopting this collective perspective can reduce the third-person impact. Conversely, when people feel detached or independent from their social group, the third-person effect is either maintained or amplified (Cro-Tool, 2022). In future studies, it would be valuable to delve deeper into the relationship between the types of hate speech and support for Facebook's regulation, considering the nuanced dynamics at play. Additionally, exploring how self-construal influences the perception of hate speech impact and support for regulation could yield insightful findings.

This study endeavours to investigate the evolving patterns of hate speech within the Malaysian population. It unveils a noteworthy correlation linking users' activity on Facebook to their awareness of hate speech. Surprisingly, the amount of time spent on Facebook does not significantly affect the frequency of encountering hate speech. Furthermore, no significant relationship is found between personal experiences of hate speech attacks and individuals' support for Facebook's regulatory measures. The study's findings serve as a fundamental reference for comprehending the intricate web of associations and relationships between these variables. However, it is vital to note that these results only establish statistical associations, and they do not imply causation. In essence, while they demonstrate a connection between the variables, they do not provide insights into the direction of this relationship or suggest causative factors. This study presents a significant limitation in its inherent focus on the specific Malaysian context. It is crucial to acknowledge the potential divergence in instances of hate speech exposure and regulatory attitudes within different cultural and societal contexts. Therefore, the generalizability of our findings to alternative social and cultural environments should be approached with caution. To gain a more comprehensive understanding of this issue, further research conducted in diverse social settings is warranted. As highlighted by Carlson and Dreher (2018) and Raynauld et al. (2018), as cited in Fernández and Farkas (2021), there is a growing body of new research that explores both the opportunities and challenges associated with the use of social media by Indigenous communities. Expanding our investigation to include such perspectives can offer valuable insights into this dynamic landscape.

It is imperative to propose that future investigations undertake a rigorous examination and underscore the significance of conscientiously addressing and managing variables of import, such as one's level of engagement on the Facebook platform, the temporal extent dedicated to Facebook usage, and the frequency of exposure to instances of hate speech within the Facebook milieu. By incorporating these variables into the research framework, a heightened degree of precision and validity can be achieved in the outcomes, particularly within the context of more expansive and comprehensive research endeavours. Prospects for subsequent research involve the possibility of conducting comprehensive in-depth interviews or employing qualitative data collection methodologies, both of which hold the potential to yield more nuanced and profound insights into users' perspectives on hate speech. These methods could shed light on their adaptive strategies for managing such content and their stances regarding regulatory measures. Integrating qualitative and quantitative research approaches presents an opportunity to achieve a more comprehensive and multifaceted comprehension of the hate speech phenomenon as it manifests on the Facebook platform.

■6.0 CONCLUSION

In conclusion, this study has shed light on various aspects of hate speech on Facebook and its broader implications. The findings have debunked the notion that hate speech on this platform is limited to racial content, highlighting the presence of other forms of hate speech. Moreover, it has unveiled a significant association between user activity on Facebook and their awareness of hate speech, while failing to establish a noteworthy correlation between time spent on the platform and the frequency of encountering hate speech. Additionally, experiences of hate speech attacks did not significantly influence support for Facebook regulation. These insights are invaluable for both Facebook and government agencies, serving as a crucial guide to proactively address the issue and foster a safer, more inclusive online community. As Hassandarvish (2020) and Jo-Ann (2019) argued, it's imperative to protect vulnerable groups, as powerful figures' remarks can inflict greater societal harm. Furthermore, as Chetty and Alathur (2018) contend, hate speech is a form of terrorism, and it is essential to have policies that counter both efficiently. Hate speech, often used as a shortcut to instant popularity, threatens free speech and fundamental human rights. Notably, hate speech is particularly pernicious when it targets traditional symbols or characteristics, such as race, religion, or gender (Chetty and Alathur, 2018). The European Court of Human Rights defines hate speech comprehensively (Council of Europe, 2022). Platforms like TikTok and YouTube have implemented strict policies against hate speech to foster inclusive communities (TikTok, 2022; Hate Speech Policy - YouTube Help, 2019). In light of these findings and perspectives, addressing hate speech is not just a matter of regulation but a collective responsibility to uphold the principles of inclusivity, respect, and freedom of expression in the digital age.

Acknowledgement

We express our gratitude to the Department of Journalism & Communication Studies within the Faculty of Humanities & Social Sciences at Southern University College, Johor, Malaysia. Furthermore, our sincere appreciation extends to the esteemed Panel reviewers from Universiti Teknologi Malaysia, whose professional guidance and valuable inputs significantly contributed to the realization of this article.

References

- Age Standard. (2014). Australian Bureau of Statistics. Retrieved from: <https://www.abs.gov.au/statistics/standards/age-standard/latest-release> Retrieved date: 6 Jan 2023
- Ahmad Tajuddin, M. S. (2020, September 9). Online Hate Speech In Malaysia - by Ahmad Tajuddin Bin Mohd Said. <https://midas.mod.gov.my/gallery/publication/midas-commentaries/213-haze-managing-another-disaster-during-covid-19-by-lt-kol-dr-maimunah-omar-2> Retrieved date: 23 Dec 2022
- Anderson, L., & Barnes, M. (2022). *Hate Speech* (Stanford Encyclopedia of Philosophy). Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/hate-speech/#WhatHateSpeech> Retrieved date: 16 Feb 2023
- Ariffin, M. A. M., Nor, M. W. M., & Razak, A. A. (2021). Cyberbullying, Its Psychological Impact and the Malaysian Legal Framework: An Analysis. *Indonesian Journal of International and Comparative Law*, 8(2), 197-214.
- Bennekom, F.V. (2018). Survey Statistical Confidence: How Many is enough? Great Brook Consulting. Retrieved from: <https://greatbrook.com/survey-statistical-confidence-how-many-is-enough/> Retrieved date: 10 Mac 2023
- Chetty, N. and Alathur, S. (2018) Hate speech review in the context of online social networks. *Aggression and Violent Behavior*, 40, 108-118. Elsevier. DOI: <https://doi.org/10.1016/j.avb.2018.05.003> Retrieved from: <https://www.sciencedirect.com/science/article/abs/pii/S1359178917301064>
- Cohen-Almagor, R. (2015). *Confronting the Internet's dark side: Moral and social responsibility on the free highway*. Cambridge University Press. <https://doi.org/10.1017/CBO9781316286649>
- Communications and Multimedia Act (1998) Act 588, Incorporating All Amendments Up To 1 January 2006. (URL)Retrieved from: https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/Act588bi_3.pdf Retrieved date: 6 Mac 2023
- Council of Europe. (2022). *Council of Europe's work on Hate Speech*. No Hate Speech Youth Campaign. <https://www.coe.int/en/web/no-hate-campaign/coe-work-on-hate-speech> Retrieved date: 26 Feb 2023
- Cro-Tool, C. (2022). *The Third-Person Effect Detailed Explanation*. <https://cro-tool.com/theories/third-person-effect> Retrieved date: 8 Feb 2023
- Fernández, A.M. and Johan Farkas, J. (2021). Racism, Hate Speech, and Social Media: A Systematic Review and Critique. *Television & New Media*, 22(2), 205-224. DOI: <https://journals.sagepub.com/doi/epub/10.1177/1527476420982230>
- Guiora, A. N., & Park, T. (2017). Free speech vs. hate speech: Exploring the legal and psychological contours. *International Journal of Law, Crime and Justice*, 49, 39-47. <https://doi.org/10.1016/j.ijlcj.2017.01.004>
- Guiora, A. Z., & Park, J. H. (2017). *Freedom of Speech and Hate Speech: Some Historical, Legal, and Philosophical Perspectives*. In W. S. Brown, S. Gaeta, & T. E. Lewis (Eds.), *Hate Crime and Hate Speech in America*. 81-92. ABC-CLIO.
- Guo, L., & Johnson, N. A. (2020). Facebook users' attitudes towards hate speech and censorship: An investigation of perceived effects, paternalism, and gender differences. *Journal of Social and Political Psychology*, 8(2), 480-497.
- Hangartner et al. (2021). Empathy-based counterspeech can reduce racist hate speech in a social media field experiment. *PNAS*, 118(50), e2116310118. DOI: <https://www.pnas.org/doi/epdf/10.1073/pnas.2116310118>
- Hassandarvish, M. (2020). Cyber racism and Covid-19: Expert weighs in on hate speech in Malaysia. *Malay Mail*. <https://www.malaymail.com/news/life/2020/05/29/cyber-racism-and-covid-19-expert-weighs-in-on-hate-speech-in-malaysia/1870481> Retrieved date: 6 Jan 2023
- Hate speech policy - YouTube Help. (2019, June). <https://support.google.com/youtube/answer/2801939?hl=en>
- Harriman, N., Shortland, N., Su, M., Cote, T., Testa, M. A., & Savoia, E. (2020). Youth Exposure to Hate in the Online Space: An Exploratory Analysis. *International Journal of Environmental Research and Public Health*, 17(22), 8531. <https://doi.org/10.3390/ijerph17228531>
- Jo-Ann, D. S. A. (2019, January 17). Does Malaysia need hate speech laws? *Malaysiakini*. <https://www.malaysiakini.com/news/460524> Retrieved date: 8 Feb 2023
- Justitia. (2020, December 17). Global Handbook on Hate Speech Laws. *The Future of Free Speech*. <https://futurefreespeech.com/global-handbook-on-hate-speech-laws/#:%7E:text=Hate%20speech%20is%20defined%20as,a%20major%20violent%20event%20and> Retrieved date: 10 Mac 2023
- Karl. (2023). The 15 Biggest Social Media Sites and Apps [2021]. *Dreamgrow*. <https://www.dreamgrow.com/top-15-most-popular-social-networking-sites/> Retrieved date: 25 Mac 2023
- Kormelink, T. G., & Meijer, I. C. (2019). A User Perspective on Time Spent: Temporal Experiences of Everyday News Use. *Journalism Studies*, 21(2), 271-286. <https://doi.org/10.1080/1461670x.2019.1639538>
- Law, C. W. (2013). Freedom of speech in Malaysia: A critical overview. *Commonwealth & Comparative Politics*, 51(1), 42-61. <https://doi.org/10.1080/14662043.2013.755228>
- Laws of Malaysia Online Version of Updated Text of Reprint Act 574 Penal Code As At 1 February 2018. (URL)Retrieved from: <https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/61339/117909/F1085941047/MYS61339%202015.pdf> Retrieved date: 12 Jan 2023
- Lua, A. (2022, March 16). Top Social Media Sites to Consider for Your Brand in 2022. *Buffer Library*. <https://buffer.com/library/social-media-sites/>
- Mia, M. N., Roy, R. B., & Rahman, M. M. (2021). Freedom of expression and hate speech: A comparative study between Bangladesh and Malaysia. *Journal of Asian and African Social Science and Humanities*, 7(2), 51-63.
- Malaysian Communications and Multimedia Commission (MCMC). (2022). Over 1,700 hate speech complaints. *Malaysian Communications and Multimedia Commission (MCMC) | Suruhanjaya Komunikasi Dan Multimedia Malaysia (SKMM)*. <https://www.mcmc.gov.my/en/media/press-clippings/over-1-700-hate-speech-complaints/> Retrieved date: 12 Jan 2023
- NapoleonCat. (2022). Facebook users in Malaysia - January 2022. <https://napoleoncat.com/stats/facebook-users-in-malaysia/2022/01/>
- Nor, M. R. M., & Ab Razak, N. H. (2017). Hate speech and the regulation in Malaysia: A legal analysis. *International Journal of Law, Government and Communication*, 2(5), 59-69. <https://doi.org/10.18178/ijlgc.2.5.59-69>
- Rosley, N. (2022). *Digital Marketing 101: Social Media Users in Malaysia*. Kode Digital Sdn Bhd. <https://kodedigital.expert/social-media-users-in-malaysia/>
- Ruwandika, T. (2018). Understanding hate speech: Challenges and approaches to defining, identifying, and countering hate speech in Sri Lanka. *Journal of Hate Studies*, 14(1), 183-214.
- Sert, E., & Başkale, H. (2022). The effect of pandemic period on social media addiction levels of university students. *Journal of Social Sciences Research*, 8(1), 24-34. <https://doi.org/10.32286/jssr.81033>
- Statista. (2022, December 5). Social media users as a percentage of the total population Malaysia 2022. Retrieved from: <https://www.statista.com/statistics/883712/malaysia-social-media-penetration/> Retrieved date: 12 Mac 2023
- TikTok. (2022, October). *Community Guidelines*. <https://www.tiktok.com/community-guidelines>
- United Nations. (n.d.). *What is hate speech?* <https://www.un.org/en/hate-speech/understandinghate-speech/what-is-hate-speech> Retrieved date: 10 Mac 2023
- Visibility, S. (2016). What Does It Mean to Be Active on Social Media — And Why Is It Important? *Seek Visibility*. <http://www.seekvisibility.com/2016/05/active-on-social-media/>
- Xu, Y. (2021). Cyberbullying and Hate Speech in Social Media. In *Handbook of Research on Social Media Applications in the Tourism and Hospitality Industry*. 45-56. IGI Global.