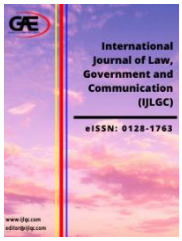
International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001



INTERNATIONAL JOURNAL OF LAW, GOVERNMENT AND COMMUNICATION (IJLGC)
www.ijlgc.com

HATE SPEECH IN THE DIGITAL AGE: A STUDY IN TERMS OF IMPACT AND SOCIAL IMPLICATIONS

Mohammad Nurhafiz Hassim[1*], Nur Nasliza Arina Mohamad Nasir[2*], Norena Abdul Karim Zamri[3]

[1] Faculty of Communication and Media Studies, Universiti Teknologi MARA, Malaysia
Email: hafiszhassim@uitm.edu.my
[2] Faculty of Communication and Media Studies, Universiti Teknologi MARA, Malaysia
Email: nasliza@uitm.edu.my
[3] Institute of Malay and Civilization (ATMA), Universiti Kebangsaan Malaysia (UKM)
Email: norena@ukm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 22.09.2024
Revised date: 10.10.2024
Accepted date: 03.11.2024
Published date: 12.12.2024

To cite this document:

Hassim, M. H., Nasir, N. N. A. M., & Zamri, N. A. K. (2024). Hate Speech In The Digital Age: A Study In Terms Of Impact And Social Implications. *International Journal of Law, Government and Communication, 9* (38), 01-12.

DOI: 10.35631/IJLGC.938001

Abstract:

Hate speech has become one of the most common and critical manifestations of social activity in the digital age. Since digital platforms have now emerged as the primary means of communication across the world, they have also appeared as the primary places that help in the propagation of hate speech. The objective of this research is to discover the impact of hate speech, and social implications based on the previous study. The method employed in this study is secondary data collected from the academic databases of Scopus, Web of Science (WOS), and Google Scholar. Data analysis in the study was done narratively, which is a type of qualitative analysis that involves interpreting the narratives. The findings of the study revealed that there are several impacts, such as the architectural characteristics of social media platforms leading to the fast dissemination of detrimental content, anonymity and fast sharing of detrimental content, and psychological consequences. Social implications discovered social categorisation and excluded vulnerable groups, difficulty in avoiding harassment, othering, prejudice becoming the new norm, universality and legal issues, impact on trust and social cohesion, and political environment and prejudice becoming the new norm. Therefore, this study aims at offering a broad understanding of the different factors that shape hate speech. It suggests future research, policymaking and practical solutions for the improvement of the digital environment and the reduction of cyberbullying.

Keywords:

Hate Speech, Digital Age, Social Media Platforms, Impact, Social Implications

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

Introduction

The advancement in technology has brought about a change in the communication process in the society through the provision of an opportunity for communication in the society as well as the sharing of information. But this technological advancement has also enabled the fast dissemination and escalation of hate speech, which is a problem for society. Hate speech, on the other hand, is a deliberate communication that is meant to offend people based on their group identity, including their race, colour, gender, or religion, and the intention is to demean them (Castellanos et al., 2023). In the context of online platforms, hate speech can be defined as any comment, post, picture, or video that is meant to degrade, offend, or even call for harm on a person or group of people based on their characteristics (Akinsanya, 2024). Hate speech may be in the form of spoken words, written words, or even through gestures and signs (Davani et al., 2023). Another trend is the growth of hate speech, which employs social networks to send obscene remarks and obscene behaviour, thus degrading the general user experience and making users susceptible to harassment (Mollas et al., 2022). Another is the fact that, in most cases, the offenders can easily create fake accounts and use them to perpetrate the acts without being easily apprehended. In order to solve this problem, most social media companies have provided the reporting systems and moderation tools for users to report and fight against hate speech and, therefore, to build a more secure online environment for all users (Alkiviadou, 2019). In addition, there are also other platforms that have incorporated the use of AI algorithms to try and prevent hate speech before it goes viral, illustrating a technological solution for this emerging problem (Garland et al., 2020).

In consideration of the impact of hate speech on people and groups, relevance to the community is critical. Cyberbullying, abuse, and other forms of hate speech are realities of contemporary society and pose a threat to society (Siegel, 2020). Such knowledge is important in order to shape the most effective interventions and policies that can avoid or reduce these detrimental effects on people and groups. Due to the rise of hate speech, especially on social media platforms, it becomes relevant to study the subject. Hate speech has also risen with the advancement in digital communication, and this is why there is a need to do more studies to capture all the dimensions of hate speech. Academic and professional people have the moral responsibility to fight against hate speech to make societies better and more tolerant. This line of research is useful to society and public safety by giving the knowledge that policymakers require in formulating laws and regulations that tackle the issue without violating the First Amendment.

Therefore, it is important to understand the social impact on hate speech for mental health, social inclusion, building the positive narrative, and for legal and ethical concerns. It is important for the development of adequate policies, technologies, and educational tools that could contribute to the formation of a society that is more tolerant and less violent. The importance of hate speech cannot be overstressed because it is crucial in the elimination of prejudice and the protection of the rights of all persons.

Literature Review

The Evolution of Hate Speech

The history of hate speech can be traced back too many years ago, and it has gone through many changes. The idea of hate speech can be dated back to prehistoric man when one person or a group of people attacked another person or a group of people because of their race, religion,

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

ethnic background or other aspects (Alkiviadou, 2018). Over the years, hate speech has been employed as a way of instilling fear, encouraging violence, and dominating the minorities. In the United States, hate speech has been a sensitive subject of debate right from the formation of the country. The First Amendment of the Constitution of the United States of America guarantees freedom of speech including hate speech which has raised questions on whether hate speech should be protected since it may offend or even harm a person (Carlson, 2023; Pomeranz et al., 2023). This was especially realised during the Civil Rights Movement of the 1960s, when people became more conscious of the effects of hate speech on vulnerable groups (Kaplan & Inguanzo, 2020).

Hate speech has been a prevalent problem in the past, but the problem has become more significant in the recent past because of the increase in the number of people engaging in online conversations. This has been researched widely, and it was discovered that hate speech aims at vulnerable or minority groups and results in social segmentation and violence (Izquierdo Montero et al., 2022). The correlation between the online discourse and the physical hate crimes has been established and it has been found that major events that affect the minorities can prompt hate speech and subsequently hate crimes against the Black and the LGBTQ+ people (Bozhidarova et al., 2023).

The task of moderation of hate speech is further complicated by the sheer volume of user-generated content, which requires the use of sophisticated machine learning techniques for detection and categorisation (Kumar et al., 2024; Mittal et al., 2023). Studies have revealed that hate speech not only challenges the social cohesion of multicultural societies in terms of intercultural communication but also intensifies the "othering" narrative in the digital space, as revealed in the cross-cultural studies, including those conducted in Malaysia (Zamri et al., 2023). It is important to understand that hate speech has been widely associated with the development of technologies and the popularity of social networks. As the internet and social networking sites become more prominent, hate speech has also emerged as a new way of expressing its hate messages. In addition, the use of the internet to disseminate the information has made it possible for people to come up with discriminating statements without the fear of being disciplined immediately (Wang, 2018). This anonymity together with the feature of being able to reach people from all over the world in seconds has enhanced the spread and influence of hate speech in our current society (Von Essen & Jansson, 2018).

However, with the recent advancement in technology, especially the social media platforms, there has been an increase in the use of hate speech. There have been debates on censorship, freedom of speech and the part played by technology firms in moderating content following pressure to deal with hate speech (Bakalis & Hornle, 2021). In conclusion, the historical analysis of hate speech shows that the attempts to regulate hate speech have always been an attempt to regulate freedom of speech while trying to protect people from harm and discrimination.

Digital Hate Speech in the Malaysian Context

With the advancement of technology and the use of social media platforms like X, Facebook, Instagram, and TikTok, hate speech has increasingly been spreading through the internet and rising around the world. Due to its unprecedented ability to spread quickly over the globe, preventing and countering digital hate speech poses particular challenges. According to the United Nations (n.d.), under international human rights law, there is no universal definition of

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

hate speech. The standards used to define it differ throughout the nations that have laws prohibiting it. Despite the difficulty in defining hate speech, some nations have addressed it through their legal systems, considering the social norms and historical backgrounds unique to each of them.

In Malaysia, where people of multiple religions, races, and cultures live together, it is very important to control hate speech from spreading in order to maintain and safeguard the country's harmony and stability. When it comes to combating digital hate speech, the Malaysian Communications and Multimedia Commission (MCMC) takes its responsibility seriously. 3,419 complaints about hate speech were handled by MCMC between October 2020 and October 2023, highlighting the growing difficulties regulatory organisations confront in halting the spread of harmful and discriminatory online conversations. The growing volume of complaints suggests that people are becoming more sensitive to the effects of hate speech. In order to resolve complaints thoroughly, MCMC conducts investigations in accordance with Section 233 of the Communications and Multimedia Act 1998 in order to identify offenders and impose sanctions on the people or organisations in charge of spreading hate speech. In addition, MCMC enforced administrative actions, including removing content and accounts, restricting access to the website, and creating abuse complaints (Aingaran, 2023).

Other than MCMC, there are various other parties in Malaysia who are serious about combating digital hate speech. Among them are The Communications and Multimedia Content Forum of Malaysia and The Centre, which collaborate to spread awareness about the importance of self-regulation and the necessity of effectively combating digital hate speech in this country. Based on data gathered by The Centre (2022) using a prototype of Artificial Intelligence (Al) named #TrackerHate (#TrackerBenci), the number of hateful tweets has increased over the past three months. In March 2022, #TrackerHate (#TrackerBenci) recorded 2,740 tweets identified as hateful and this figure increased to 3,088 in April 2022 and decreased slightly to 3,004 in May 2022. Furthermore, the tracking system detected 34% of the terms and/or phrases as possibly hateful. The data was recorded exclusively on Twitter, where the #HateTracker was programmed to identify the number of hate tweets by week (Forum Kandungan Komunikasi dan Multimedia Malaysia, 2022).

Many studies have shown significant trends in hate speech touching on issues of race, religion, gender, ethnicity and sexuality in Asian and European contexts (Husni, 2019; Kang et al., 2020; Kim-Wachutka, 2020; Morada, 2021; Wan Mohd nor & Gale, 2021; Bayer & Petra, 2020; Bleich, 2017; Howard, 2017). Similar to this, hate speech in Malaysia also often touches on issues related to race, religion, gender, ethnicity, and sexuality according to the study of Zamri et al. (2023). In addition, the Malaysian government is serious about dealing with hate speech that touches on issues related to the 3Rs (race, religion and royalty). As reported by Sinar harian (2024), MCMC informed that a total of 2,004 contents involving hateful speech and touching aspects of race, religion and royalty (3Rs) that have the potential to trigger violence and discrimination have been taken down since January 2023 until March 2024. MCMC will not compromise and tolerate any spread of extreme provocative content on social media. In addition, MCMC is working with the Royal Malaysian Police (PDRM) to track and trace the account owners involved and bring them to justice.

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

Addressing and combating hate speech is important. It requires a comprehensive strategy that mobilises the entire community. All people and institutions, including the public and commercial sectors, the media, Internet companies, religious leaders, educators, young people, and civil society, have a moral obligation to strongly condemn hate speech and play an important role in combating this evil phenomenon (United Nation, n.d.). To stop hate speech from spreading and dividing Malaysian society, all relevant parties—including the government, academics, researchers, authorities, and the community—should work together to combat it, irrespective of a person's ethnicity, religion, or cultural background.

Methodology

This conceptual paper employs a secondary data analysis approach and thus, the study entails articles published from 2019 to 2023 from peer-reviewed journals. To complement the results a narrative analysis is performed to determine the impact of hate speech and social implications in the digital age. The data for this paper was collected from academic databases including Google Scholar, Scopus, and Web of Science (WoS). These databases were selected because they offer a large number of academic articles and peer-reviewed journals on hate speech. The search terms employed were ‘hate speech’, social implications, ‘digital age’, ‘online hate speech’, and ‘social media’. The articles were searched with the intention of retrieving only those that were published between 2019 and 2023 to reflect the current literature in the field. The identified literature was grouped into themes and topics that arose from the analysis of the data collected. This way of structuring the data was useful as it provided a clear and coherent structure to the explanation of the various features of hate speech and its social environment. Apart from the thematic analysis, a narrative synthesis was also conducted to review the findings as presented in the selected literature. It involved integrating the results of the studies summarising the results, comparing the results, and relating the results of the studies. This approach was useful in capturing the nature of the topic and the fact that hate speech is not a one-dimensional issue but has many dimensions due to the current use of social media. To improve the methodological quality of the study, only published articles in refereed journals were employed in the analysis. This criterion was employed to maintain academic integrity and ensure that the conclusion and the insights given in this paper are based on credible information. Furthermore, the use of several databases and the thematic analysis of the data helped to reduce the potential bias and to get an overall view of the current state of the research. Therefore, since this is a conceptual paper, secondary data analysis and narrative review research methods offer a fair and inclusive method of presenting the social impacts of hate speech from the current and relevant literature.

Discussion

The Impact of Digital Platforms on Hate Speech Proliferation

The phenomenon of hate speech in the context of social media is one of the most important and significant challenges of the modern digital society. With the intensification of the use of social networks and other online resources in everyday communication, they have also become a source of dissemination of negative information, including hate speech. The availability of social media platforms has accelerated the diffusion of hate speech, which is usually initiated by marginal groups and then extended to other platforms, as demonstrated by the research on Reddit, where people’s participation in hate speech subreddits resulted in elevated hate speech in other areas of the site (Schmitz et al., 2022). This is compounded by the technological characteristics of social media platforms through which hate speech is conducted, which are

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

both enablers of hate speech and possibly its inhibitors. For example, features like anonymity and the ability to share content quickly are the reasons for the use of hate speech, whereas technology-based solutions could be a prevention measure (Weber et al., 2023). Benigni et al. (2019) note that the architecture of social media, especially the recommendation systems, the interaction dynamics, and the content sharing systems, are instrumental in the spread of hate speech. These features are meant to encourage users to stick around and engage more, but they also make it easier for hate speech to reach more people in a shorter amount of time, thus worsening the effects.

In addition, the social media and technological connectivity of the world makes hate speech more impactful and provides anonymity and the ability to spread hate speech. In many social networks, people can be anonymous, which means that some people can easily post something that will offend others without getting a real-life punishment for this. This can lead to aggressive and abusive behaviour that users would not otherwise exhibit if their identity was not concealed (Siegel, 2020). In addition, anonymity can lead to a lack of trust within online communities since people feel endangered and vulnerable, especially when they cannot know who stands behind negative comments. This can make the online environment toxic, and there is no healthy discussion because of the presence of hate speech (Ascher & Umoja Noble, 2019). As noted by Ghufron et al. (2024); Hassim & Ahmad (2023), this behaviour is made worse by the fact that most social media platforms allow anonymity and can fuel large-scale conflict and harassment.

Apart from that, hate speech contributes to stress, anxiety, and perceived vulnerability among the target persons and groups. This is especially so for the minority groups, who are usually the victims of hate speech. These effects are further compounded by the fact that one can be targeted online and, in many cases, on a global scale, with little to no protection or backup (Stahel & Baier, 2023; Zamri et al., 2023). In a social aspect, hate speech on the internet leads to the segregation of societies. It promotes hostility and social fragmentation and erodes social solidarity and reliance. This can be seen in how hate speech can cause actual harm, such as hate crimes, and social unrest in a society. Social media facilitates the dissemination of toxic ideas, which can lead to the radicalisation of people and groups and deepen social cleavages (Buturoiu & Corbu, 2020). In politics, hate speech can help shape the direction of the discourse and the policy. It can distort people's perception and discussion, thus resulting in the exclusion of some groups in political affairs. The era of the internet and social networks has brought such phenomena as echo chambers, in which people are provided only with information that supports their biases, including hate speech. This can result in the development of a skewed perception of social problems and the inability to foster positive discussions and policymaking (Salma, 2019). Furthermore, hate speech regulation has also been complicated by the digital age. Since the internet transcends boundaries, hate speech is likely to spread across borders, making it difficult to implement national laws and regulations. It is common for platforms to have conflicting objectives of free speech and safety of users and thus have an irregular approach to the policies against hate speech (Aljasir, 2023).

However, there are some that have been taken to attempt to mitigate the impacts of hate speech on the internet. These are the use of algorithms to block out negative content, the social media companies educating the users on the effects of the negative posts, and the policies that regulate the conduct of the users (Garland et al., 2022; Saha et al., 2019). These are the algorithms that can help to filter out hate speech, awareness campaigns to increase people's understanding of

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

the issue and make them less hostile, and the laws that can punish those who use hate speech. However, such measures require constant updates because of the dynamics of the digital environment and the activities of those who spread hatred (Ibrahimova, 2023; Zamri et al., 2023). Consequently, it is possible to conclude that hate speech in the context of the digital age has severe and multiple consequences that are reflected in the sphere of people's psychological and social conditions, social interactions, and political processes. Minimising these impacts thus requires a technological, educational, and legal approach. As the usage of digital platforms is expanding, it is necessary to increase the effectiveness of combating hate speech and its consequences.

Social Implications of Hate Speech in the Digital Age

The advancement in technology, especially in the use of the internet has greatly enhanced the communication system and also enhanced the spread of hate speech. The social implications of hate speech in this context are therefore manifold at the individual, group and structural levels. Firstly, the use of hate speech on online platforms contribute to the intensification of social fragmentation and the exclusion of minorities. The openness and freedom of speech that social media offers enable the spread of prejudice in a short span of time, thus causing heightened discrimination, and social division. This is especially worrisome as it may lead to the creation of settings in which hate speech is acceptable and people's isolation is deepened (Buturoiu & Corbu, 2020). Furthermore, hate speech affects the targeted individuals psychologically in a very significant way. Stress, anxiety, and depression are some of the effects of sexual violence and may result in chronic mental health disorders. The fact is that people cannot always avoid such speech, which increases its psychological impact on the subject (Stahel & Baier, 2023; Zamri et al., 2023).

These studies show that there are negative impacts of hate speech on victims and that they experience poor mental health. In a study carried out in Switzerland, the authors noted that people who are on the receiving end of hate speech online are more insecure not only when they are online but also in their real lives (Paz et al., 2020). Also, given that hate speech is constantly available on the internet and can remain visible for an indefinite period, the victimisation emotions are prolonged. The victims are unable to avoid the digital harassment as the smartphones and the ever-connected nature of the world ensure that they are constantly exposed to the messages that harm them, making them feel that their personal spaces are being intruded upon (Paz et al., 2020).

Subsequently, hate speech on social media promotes prejudice and discrimination which may extend to other aspects of life. Research has established that hate speech leads to the socialisation of prejudice and therefore increases the tolerance level of society to bigotry (Castaño-Pulgarín et al., 2021). This environment not only affects the targeted groups but also poses a threat to society's cohesiveness as it fosters division and hostility between groups. The concept of 'othering' is closely associated with hate speech and is discussed in detail in the context of social media, which has emerged as a major enabler of such ideologies. This amplification can result in real life outcomes such as violence and ethnic cleansing. Analysis of different studies shows that this problem is rather complex and has many sides. Social media sites are inherently built to share information quickly, and this includes hate speech against people of colour, religious beliefs, or gender. The algorithms that govern these platforms are usually set to increase engagement, which in turn, fuels the circulation of material that is provocative or bigoted, thereby fuelling the spread of hatred. This is rather worrying since such

International Journal of Law, Government and Communication
EISSN : 0128-1763
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

attitudes can become acceptable and promote exclusion of targeted groups, a phenomenon referred to as 'othering' (Alorainy et al., 2019; Zamri et al., 2023). The digital space thus becomes a breeding ground for such extremists to give a boost to the existing societal cleavages and may culminate in violent actions.

From the research evidence, it is clear that hate speech in the digital age does pose significant threats to the democratic system of governance. As a result of the use of social networks, hate speech has appeared and has spread, which poses a danger to democracy and society. First, hate speech is much more significant in the context of the digital world because it spreads much faster and affects people from different countries. This makes it hard to moderate because the laws of different countries and cultural acceptability of free speech and speech that is considered hate speech are not the same (Topidi, 2022). This is because hate speech is fast and when spread causes division of society and erodes trust in democracy since it targets vulnerable groups and destabilizes, (Assen, 2023). But the political environment is also to blame for the upsurge of hate speech. The media has been blamed for echoing hate speech from politicians and this has led to enhanced use of hate speech on social media and in society. This is a threat to not only individual freedoms but also democracy as hate speech isolates the minorities and makes people less engaged in the citizenship (Laub, 2019).

Therefore, the social implication of hate speech in the era of social networking services is diverse and complex, involving the mental health of the recipients of hate speech, social participation, and democracy. To address these concerns, there must be a multi-faceted approach that includes regulation of such sites but with consideration of the first amendment rights, raising user awareness and making them prepared. This paper calls for collaboration between governments, technology companies, and civil society organisations to address the negative effects of hate speech in the online community.

Conclusion

As hate speech has proliferated in the digital age on social media platforms, researchers and policy makers alike are in urgent need to both understand its extent and determine what can be done about it. Although recent research has illuminated what hate speech is and what its consequences might be, there are still many unanswered questions. Future research should move beyond a technological lens to better understand the technological, psychological and social dynamics of hate speech and should focus on the development of effective interventions to regulate hate speech.

However, one of the most important future works should be the development of more sophisticated technological tools that help to detect and moderate hate speech. Social media platforms are changing, and its usage is growing and there are not enough methods to moderate current modern day, and ideologists are manual with basic wood algorithm that cannot update with the rated and fined pace of the internet. Specifically, research should be directed on developing more advanced machine learning algorithms and helping the AI systems to detect more subtle kinds of hate speech which may bypass the traditional filters devised by programmers, like coded jargon or forms of discrimination that are barely noticed.

Secondly, it is necessary to investigate the ethical consequences of the deployment of AI based solutions, especially focusing on a content moderation balance and its necessity to a free speech preservation. Automated systems are very good at identifying harmful content, but they can

Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

also suppress legitimate expressions of opinion. Future work can increase our understanding of how AI systems can help balance freedom of expression with the filtering of hate speech where free speech and harmful speech overlap. The study emphasises the impact and social implications of social media in the spread of hate speech. As much as these platforms have opened space for free information flow and communication, they have also opened up space for hate speech. The evaluation of the selected articles shows that technology, user behaviour, and platform policies work in a rather intricate manner that can either reduce or increase the presence of hate speech. In conclusion, the study reiterates the need to enhance the content moderation approaches, enhance the enforcement of the policies of the platforms, and create awareness of the effects of hate speech. In addition, the study raises the need for more research with the aim of enhancing knowledge on the dynamics of hate speech in the social media age and the most appropriate ways of tackling it.

Acknowledgment

The work was supported by the Ministry of Higher Education (MoHE) of Malaysia through the Fundamental Research Grant Scheme (Ref: FRGS/1/2021/SS0/UITM/02/15).

References

Aingaran, R. (2023, November 22). Over 3,400 hate speech complaints processed: MCMC. The Sun. https://thesun.my/local-news/over-3400-hate-speech-complaints-processed-mcmc-OA11778521#google_vignette

Akinsanya, T. (2024). Media And Hate Speech: A Discursive Study of Hate Speech On Nairaland FORUM. Sprin Journal of Arts, Humanities and Social Sciences, 3(4), 52–61. https://doi.org/10.55559/sjahss.v3i4.268

Alkiviadou, N. (2018). The Legal Regulation of Hate Speech: The International and European Frameworks. Politička Misao, 55(4), 203–229. https://doi.org/10.20901/pm.55.4.08

Alkiviadou, N. (2019). Hate speech on social media networks: Towards a regulatory framework? Information & Communications Technology Law, 28(1), 19–35. https://doi.org/10.1080/13600834.2018.1494417

Aljasir, S. (2023). Effect of online civic intervention and online disinhibition on online hate speech among digital media users. Online Journal of Communication and Media Technologies, 13(4), e202344. https://doi.org/10.30935/ojcmt/13478

Alorainy, W., Burnap, P., Liu, H., & Williams, M. L. (2019). "The Enemy Among Us": Detecting Cyber Hate Speech with Threats-based Othering Language Embeddings. ACM Transactions on the Web, 13(3), 1–26. https://doi.org/10.1145/3324997

Ascher, D. L., & Umoja Noble, S. (2019). Unmasking Hate on Twitter: Disrupting Anonymity by Tracking Trolls. In S. J. Brison & K. Gelber (Eds.), Free Speech in the Digital Age (1st ed., pp. 170–188). Oxford University PressNew York. https://doi.org/10.1093/oso/9780190883591.003.0011

Assen, M. D. (2023). Challenges of Governance in an Age of Disinformation and Hate Speech in Africa: A Lesson from the Post-2018 Ethiopian Experience. https://doi.org/10.33774/apsa-2023-ns4lj

Bakalis, C., & Hornle, J. (2021). The Role of Social Media Companies in the Regulation of Online Hate Speech. In A. Sarat (Ed.), Studies in Law, Politics, and Society (pp. 75–100). Emerald Publishing Limited. https://doi.org/10.1108/S1059-433720210000085005

Bayer, J., & Petra, B. Á. R. D. (2020). Hate speech and hate crime in the EU and the evaluation of online content regulation approaches.

International Journal of Law, Government and Communication
EISSN : 0128-1763
IJLGC
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

Benigni, M. C., Joseph, K., & Carley, K. M. (2019). Bot-ivistm: Assessing Information Manipulation in Social Media Using Network Analytics. In N. Agarwal, N. Dokoohaki, & S. Tokdemir (Eds.), Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining (pp. 19–42). Springer International Publishing. https://doi.org/10.1007/978-3-319-94105-9_2

Bleich, E. (2017). Freedom of expression versus racist hate speech: Explaining differences between high court regulations in the USA and Europe. In Marcel, M. & Ralph, G. (Eds.), Regulation of Speech in Multicultural Societies (pp. 110–127). Routledge.

Bozhidarova, M., Chang, J., Ale-Rasool, A., Liu, Y., Ma, C., Bertozzi, A. L., Brantingham, P. J., Lin, J., & Krishnagopal, S. (2023). Hate speech and hate crimes: A data-driven study of evolving discourse around marginalized groups. 2023 IEEE International Conference on Big Data (BigData), 3107–3116. https://doi.org/10.1109/BigData59044.2023.10386312

Buturoiu, D. R., & Corbu, N. (2020). Exposure to Hate Speech in the Digital Age. Effects on Stereotypes About Roma People. Journal of Media Research, 13(2(37)), 5–26. https://doi.org/10.24193/jmr.37.1

Carlson, C. R. (2023). On Shaky Ground: Reconsidering the Justifications for First Amendment Protection of Hate Speech. Communication Law and Policy, 28(2), 124–151. https://doi.org/10.1080/10811680.2023.2193571

Castaño-Pulgarín, S. A., Suárez-Betancur, N., Vega, L. M. T., & López, H. M. H. (2021). Internet, social media and online hate speech. Systematic review. Aggression and Violent Behavior, 58, 101608. https://doi.org/10.1016/j.avb.2021.101608

Castellanos, M., Wettstein, A., Wachs, S., Kansok-Dusche, J., Ballaschk, C., Krause, N., & Bilz, L. (2023). Hate speech in adolescents: A binational study on prevalence and demographic differences. Frontiers in Education, 8, 1076249. https://doi.org/10.3389/feduc.2023.1076249

Davani, A. M., Atari, M., Kennedy, B., & Dehghani, M. (2023). Hate Speech Classifiers Learn Normative Social Stereotypes. Transactions of the Association for Computational Linguistics, 11, 300–319. https://doi.org/10.1162/tacl_a_00550

Forum Kandungan Komunikasi dan Multimedia Malaysia. (2022, June 22). Ungkapan kebencian atau hate speech berleluasa di media sosial. Dagagnews.com. https://www.dagangnews.com/article/ungkapan-kebencian-atau-hate-speech-berleluasa-di-media-sosial-16517

Garland, J., Ghazi-Zahedi, K., Young, J.-G., Hébert-Dufresne, L., & Galesic, M. (2020). Countering hate on social media: Large scale classification of hate and counter speech. Proceedings of the Fourth Workshop on Online Abuse and Harms, 102–112. https://doi.org/10.18653/v1/2020.alw-1.13

Garland, J., Ghazi-Zahedi, K., Young, J.-G., Hébert-Dufresne, L., & Galesic, M. (2022). Impact and dynamics of hate and counter speech online. EPJ Data Science, 11(1), 3. https://doi.org/10.1140/epjds/s13688-021-00314-6

Ghufron, M. I., Supriyati, E., & Listyorini, T. (2024). Analisa Jejaring Sosial Terhadap Fenomena Cyberbullying Fandom K-Pop pada Sosial Media Twitter. JISKA (Jurnal Informatika Sunan Kalijaga), 9(2), 79–93. https://doi.org/10.14421/jiska.2024.9.2.79-93

Hassim, M. N., & Ahmad, N. F. (2023). Understanding Cyberbullying On Twitter Among K-Pop Fans In Malaysia. Asian People Journal (APJ), 6(2), 91–108. https://doi.org/10.37231/apj.2023.6.2.537

Howard, E. (2017). Freedom of expression and religious hate speech in Europe. Routledge.

International Journal of Law, Government and Communication
EISSN : 0128-1763
IJLGC
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001
Husni, H. (2019). Moderate muslims' views on multicultural education, freedom of expression, and social media hate speech: An empirical study in west java Indonesia. Jurnal Penelitian Pendidikan Islam, 7(2), 199–224.

Ibrahimova, M. (2023). In Depth: Unveiling hate speech in the digital world. The UNESCO Courier, 2023(4), 44–45. https://doi.org/10.18356/22202293-2023-4-12

Izquierdo Montero, A., Laforgue-Bullido, N., & Abril-Hervás, D. (2022). Hate speech: A systematic review of scientific production and educational considerations. Revista Fuentes, 2(24), 222–233. https://doi.org/10.12795/revistafuentes.2022.20240

Kang, M., Lee, J., & Park, S. (2020). Meta-analysis on hate speech studies in South Korea. In Hate Speech in Asia and Europe (pp. 7–22). Routledge.

Kaplan, M. A., & Inguanzo, M. M. (2020). The Historical Facts about Hate Crime in America The Social Worker's Role in Victim Recovery and Community Restoration. Journal of Hate Studies, 16(1). https://doi.org/10.33972/jhs.147

Kim-Wachutka, J. J. (2020). Hate speech in Japan: Patriotic women, nation and love of country. In Hate Speech in Asia and Europe (pp. 23–42). Routledge.

Kumar, S., Musharaf, D., & Bhushan, B. (2024). Taming The Hate: Machine Learning Analysis Of Hate Speech. 2024 2nd International Conference on Disruptive Technologies (ICDT), 269–273. https://doi.org/10.1109/ICDT61202.2024.10489362

Laub, Z. (2019, June 7). Hate Speech on Social Media: Global Comparisons. Council on Foreign Relations. https://www.cfr.org/backgrounder/hate-speech-social-media-global-comparisons

Mittal, H., Chauhan, K. S., & Shambharkar, P. G. (2023). Study on Optimizing Feature Selection in Hate Speech Using Evolutionary Algorithms. In P. Dutta, S. Chakrabarti, A. Bhattacharya, S. Dutta, & C. Shahnaz (Eds.), Emerging Technologies in Data Mining and Information Security (Vol. 490, pp. 707–720). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-4052-1_70

Mollas, I., Chrysopoulou, Z., Karlos, S., & Tsoumakas, G. (2022). ETHOS: A multi-label hate speech detection dataset. Complex & Intelligent Systems, 8(6), 4663–4678. https://doi.org/10.1007/s40747-021-00608-2

Morada, N. M. (2021). Myanmar. Journal of International Peacekeeping, 24(3–4), 428–466.

Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate Speech: A Systematized Review. SAGE Open, 10(4), 215824402097302. https://doi.org/10.1177/2158244020973022

Pomeranz, J. L., Merrill, T. G., & Schroth, K. R. J. (2023). The First Amendment. In J. L. Pomeranz, T. G. Merrill, & K. R. J. Schroth, Public Health Law in Practice (1st ed., pp. 106–162). Oxford University PressNew York. https://doi.org/10.1093/oso/9780197528501.003.0004

Saha, K., Chandrasekharan, E., & De Choudhury, M. (2019). Prevalence and Psychological Effects of Hateful Speech in Online College Communities. Proceedings of the 10th ACM Conference on Web Science, 255–264. https://doi.org/10.1145/3292522.3326032

Salma, A. N. (2019). Defining Digital Literacy in the Age of Computational Propaganda and Hate Spin Politics. KnE Social Sciences. https://doi.org/10.18502/kss.v3i20.4945

Schmitz, M., Muric, G., & Burghardt, K. (2022). Quantifying How Hateful Communities Radicalize Online Users. 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 139–146. https://doi.org/10.1109/ASONAM55673.2022.10068644

International Journal of Law, Government and Communication
EISSN : 0128-1763
IJLGC
Volume 9 Issue 38 (December 2024) PP. 01-12
DOI 10.35631/IJLGC.938001

Siegel, A. A. (2020). Online Hate Speech. In N. Persily & J. A. Tucker (Eds.), Social Media and Democracy: The State of the Field, Prospects for Reform (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781108890960

Sinar Harian. (2024). Kandungan mudarat di media sosial meningkat ketara – SKMM. Astro Awani. https://www.astroawani.com/berita-malaysia/kandungan-mudarat-di-media-sosial-meningkat-ketara-skmm-464363

Stahel, L., & Baier, D. (2023). Digital Hate Speech Experiences Across Age Groups and Their Impact on Well-Being: A Nationally Representative Survey in Switzerland. Cyberpsychology, Behavior, and Social Networking, 26(7), 519–526. https://doi.org/10.1089/cyber.2022.0185

Topidi, K. (2022). Minority Identity in Digital Governance and the Challenges of Online Hate Speech and Content Regulation. In A.-M. Bíró & D. Newman, Minority Rights and Liberal Democratic Insecurities (1st ed., pp. 135–155). Routledge. https://doi.org/10.4324/9781003239871-10

United Nations. (n.d.). Understanding hate speech. https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech

Von Essen, E., & Jansson, J. (2018). Cyberhate anonymity and the risk of being exposed. AoIR Selected Papers of Internet Research. https://doi.org/10.5210/spir.v2018i0.10510

Wan Mohd Nor, M., & Gale, P. (2021). Growing fear of Islamization: Representation of online media in Malaysia. Journal of Muslim Minority Affairs, 41(1), 17–33.

Wang, Z. (2018). Anonymity Effects and Implications in the Virtual Environment: From Crowd to Computer-Mediated Communication. Social Networking, 07(01), 45–62. https://doi.org/10.4236/sn.2018.71004

Zamri, N. A. K., Mohamad Nasir, N. A., Hassim, M. N., & Ramli, S. M. (2023). Digital hate speech and othering: The construction of hate speech from Malaysian perspectives. Cogent Arts & Humanities, 10(1), 2229089. https://doi.org/10.1080/23311983.2023.2229089