Socio-Informational Dynamics of Mis/Disinformation in Malaysia's Digital Ecosystem: An Integrative Review

Kenneth Tze Wui Lee[1*], Zhi Xin Sau[2]

[1,2] Department of Mass Communication, Faculty of Creative Industries, Universiti Tunku Abdul Rahman, Malaysia.

The alarming surge of mis/disinformation in Malaysia reflects a complex form of information disorder shaped by digital platform dynamics, ethno-political sensitivities, and contested legal frameworks. This study examines the socio-informational dynamics that underpin the construction, diffusion, and contestation of false narratives in Malaysia's digital ecosystem. Using an integrative literature review approach, insights from 28 selected sources related to mis/information in the Malaysian context, including journal articles, books, and grey literature, are synthesized to highlight key patterns shaping the digital information landscape. The analysis draws on interdisciplinary theoretical frameworks, including Diffusion of Innovations, Intergroup Threat, and Securitization theories, to explore how technological, psychological, and political factors influence the circulation of mis/disinformation. By situating the discussion within these frameworks, the study provides a nuanced understanding of Malaysia's information environment and highlights avenues for research, policy, and community-driven initiatives. The review identifies recurring socio-informational dynamics, including the rapid digital diffusion of false content and its cognitive reinforcement, the strategic use of identity-based narratives to exert intergroup influence, and the securitization of disinformation through legal frameworks and contested public responses. These insights highlight the importance of developing localized educational programs, media literacy initiatives, and evidence-informed strategies to strengthen information integrity in Malaysia's diverse society.

Keywords: Information disorder, digital ecosystem, political disinformation, deepfakes, media and information literacy

INTRODUCTION

Advancements in technology and the widespread use of digital platforms have transformed how information is shared and consumed. However, this transformation has also facilitated the mass dissemination of false information, encompassing both misinformation and disinformation. While misinformation refers to the unintentional spread of inaccuracies, disinformation is the deliberate creation and sharing of falsehoods to deceive the public, often for social, political, or economic gain (Wardle, 2017). Recognizing both the complexity and the societal consequences of this phenomenon, scholars have moved beyond a simple "real versus fake" dichotomy, with Wardle and Derakhshan (2017) proposing the term "information disorder" to better distinguish between misinformation, disinformation, and malinformation. Their framework, which outlines seven types of information disorder, including satire, misleading content, and manipulated media, has been expanded by others who examine its intersection with propaganda and synthetic media like deepfakes (Chesney & Citron, 2019; Fallis, 2015; Tandoc et al., 2018).

The spread of mis/disinformation is not a recent phenomenon. Historically, technological shifts have consistently amplified its reach. Gutenberg's invention of the printing press in the 15th century revolutionized textual production, allowing both information and mis/disinformation to circulate with unprecedented speed and scale (Briggs & Burke, 2018; Eisenstein, 2012). This newfound accessibility, supported by the establishment of paper mills, reshaped public discourse and knowledge transmission while enabling the spread of misleading narratives. An example of this was the European witch hunts (c. 1450s-1750s), which were fueled by sensationalist rumors and religious anxieties circulated through printed pamphlets, leading to mass hysteria and wrongful persecutions (Krznaric, 2024; Levack, 2016). A similar dynamic reemerged in the late 19th century with the rise of "yellow journalism," where newspapers published exaggerated or fabricated stories to manipulate public sentiment and boost circulation as literacy rates and urban populations grew (Campbell et al., 2020).

In the contemporary digital age, the velocity and impact of mis/disinformation have intensified, particularly within political contexts. The 2016 United States presidential election served as a watershed moment, where the widespread circulation of fabricated stories highlighted how false narratives could decisively shape electoral perceptions and influence democratic processes (Allcott & Gentzkow, 2017). During this period, the term "fake news" became highly politicized, frequently used to dismiss unfavorable media coverage, further eroding public trust in information (Ross & Rivers, 2022). The global COVID-19 pandemic further magnified the crisis, prompting the World Health Organization (2020) to coin the term "infodemic" to describe the overwhelming surge of accurate and misleading information that complicated public health efforts. Across the globe, false narratives about the virus and vaccines spread rapidly, fueling public confusion, promoting hesitancy, and eroding trust in scientific and governmental institutions.

Malaysia, a multiethnic and hyper-connected society, has been affected by this global information disorder. As of January 2025, 97.7% of the nation's population was online, with 70.2% active on social media (Kemp, 2025). This extensive connectivity makes platforms like Facebook, TikTok, and Instagram central to public discourse and a fertile ground for targeted mis/disinformation. The COVID-19 lockdown period exemplified this, as increased social media usage fueled the rapid spread of falsehoods, creating public confusion and compelling government intervention ("MOH acted on 14 viral fake news," 2022). Political mis/disinformation has been a persistent challenge, especially during electoral cycles. For instance, during the 2022 general elections, narratives alleging electoral fraud, doctored images of riots, and misleading information about voting procedures circulated widely on WhatsApp and Facebook, undermining public trust in the democratic process and prompting corrective action from electoral authorities ("Fake news, online provocations dropped," 2022; "GE15: EC tells voters to ignore fake news," 2022; Lund, 2022).

What makes Malaysia's digital ecosystem particularly complex is the deep intertwining of ethnic identity, religion, and politics (Weiss, 2013). This hyper-connected, multiethnic landscape provides a uniquely potent environment for mis/disinformation, where false narratives are often designed to exploit long-standing societal sensitivities. Emotionally charged content targeting communal anxieties is amplified through platforms like WhatsApp, Facebook, and TikTok, often circulating in vernacular languages such as Bahasa Malaysia, Chinese, and Tamil, allowing it to penetrate specific ethno-cultural spheres with precision (George, 2022; Tapsell, 2023; Wan Hashridz et al., 2025). The strategic weaponization of racialized, religious, and economic narratives is a recurring issue, with disinformation campaigns fabricating stories of religious conversion, alleging ethnic favoritism, or inciting fears of political instability to shape public opinion and behavior (George, 2022; Weiss, 2013). Such dynamics reveal that disinformation does not simply mislead but actively mobilizes fears and anxieties between groups, exploiting deep-rooted social sensitivities. In this context, mis/disinformation is not merely a series of isolated falsehoods but a systemic mechanism that actively exploits and deepens existing social tensions, particularly during elections or public health crises (George, 2022).

The consequences of this unchecked diffusion of mis/disinformation are significant, posing profound challenges to Malaysia's public trust and stability. The rapid spread of false news damages trust in key national institutions, polarizes communities, and undermines democratic processes (Daud & Zulhuda, 2020; George, 2022). An example was the 2017 incident involving false allegations against Tabung Haji, the nation's pilgrimage fund, which ignited widespread public alarm and eroded confidence in a critical institution ("Rafizi's Tabung Haji claims," 2018). Beyond politics, the societal impact is pervasive. During the COVID-19 pandemic, mis/disinformation regarding vaccine side effects fueled

public hesitancy, hampering effective health interventions and forcing the Ministry of Health to actively debunk viral falsehoods (Loomba et al., 2021; "MOH acted on 14 viral fake news," 2022). Such campaigns also inflict significant psychological distress, including heightened anxiety and fear, and can disrupt economic stability by shaking consumer confidence and damaging corporate reputations (He & Sannusi, 2024; Shoib et al., 2023).

AI-generated deepfakes and algorithm-driven content curation intensify these challenges. Deepfake technology, which enables the creation of hyper-realistic fabricated videos and audio, presents a formidable threat to public perception ("The threat of deepfake," 2019). Simultaneously, social media algorithms designed to maximize user engagement can inadvertently create echo chambers, amplifying emotionally charged and polarizing mis/disinformation tailored to individual biases (Karas, 2024). These dynamics are both technological and psychological, as individuals are more inclined to accept and share content that resonates with their pre-existing beliefs, making falsehoods especially difficult to dislodge once embedded in public discourse. This algorithmic amplification has been particularly influential in shaping voter perceptions in Malaysia's complex multiethnic landscape, heightening vulnerabilities to manipulation during events like the 2022 general election (Tapsell, 2023). The growing proliferation of AI-generated deepfakes further complicates this ecosystem, necessitating a localized and integrative understanding to inform effective countermeasures (Neyazi et al., 2024; Zainudin et al., 2024).

Despite extensive global scholarship on mis/disinformation, its localized dynamics in Malaysia's pluralistic and politically contested environment remain underexplored. In this hyper-connected and ethnically diverse democracy, false narratives are often shaped by identity-based appeals rooted in ethnicity, religion, and language, and their circulation is further enabled by Malaysia's complex socio-political landscape, where institutional arrangements, historical identity politics, and fragmented media ecosystems create uneven and unpredictable vulnerabilities to mis/disinformation (Akram et al., 2023; George, 2022; Radue, 2019; Rozwan et al., 2023; Tapsell, 2023; Weiss, 2013). Moreover, emerging technological affordances, such as AI-generated deepfakes, introduce new dynamics that alter the speed, credibility, and reach of falsehoods.

Addressing these knowledge gaps is not only academically significant but also practically urgent, as the unchecked diffusion of mis/disinformation risks undermining public trust, destabilizing democratic institutions, and compromising crisis governance. Government responses often frame false information not only as a communicative challenge but as a broader threat to public order and national stability, a process scholars describe as securitization. In Malaysia, this framing has been institutionalized through several legislative and regulatory measures, where "fake news" and other perceived harmful information are constructed as threats to national security.

However, scholars argue that such securitization often faces challenges due to questions of legitimacy and public resistance (Neo, 2021). Policy interventions may remain fragmented and ineffective without context-sensitive insights, perpetuating systemic vulnerabilities.

Against this backdrop, the study aims to critically identify emerging socio-informational patterns of mis/disinformation in Malaysia through an integrative review of relevant scholarly literature. By synthesizing existing evidence, it explores the structural, technological, and socio-cultural dynamics that may underpin mis/disinformation in Malaysia's hyper-connected, multiethnic society. In doing so, it provides a foundational understanding for developing nuanced, context-specific strategies to mitigate the societal risks associated with false narratives in the digital age.

METHODOLOGY

This study employs a qualitative design with an integrative literature review method, grounded in a constructivist epistemological position and an interpretivist research paradigm. This perspective assumes that reality is socially constructed and that understanding mis/disinformation requires exploring the multiple meanings individuals ascribe to these phenomena within specific social, political, and technological contexts (Berger & Luckmann, 1966; Whittemore & Knafl, 2005). The integrative review approach is particularly suited for synthesizing diverse forms of qualitative evidence, theoretical perspectives, and conceptual insights, allowing for a comprehensive understanding of complex, context-dependent issues such as mis/disinformation in Malaysia (Snyder, 2019; Torraco, 2005; Whittemore & Knafl, 2005).

To guide this methodological approach, the study draws on an integrative theoretical orientation that conceptualizes mis/disinformation as a layered socio-informational process shaped by intersecting technological, psychological, social, and political dynamics. At the technological-psychological level, Rogers' (1983) Diffusion of Innovations and Wason's (1960) work on confirmation bias explain how false content spreads through Malaysia's digital networks and gains traction by aligning with users' existing beliefs, reinforced by algorithmic amplification and echo chambers. Stephan and Stephan's (2000) Intergroup Threat Theory highlights how disinformation exploits ethnic and religious sensitivities, intensifying perceptions of threat and polarization. At the institutional level, Wæver's (1995) Securitization Theory illustrates how the state frames mis/disinformation as a security concern to justify regulation, while such framings remain contested by civil society. These perspectives informed the coding framework and interpretive strategies, ensuring analysis accounted for technological infrastructures, cognitive tendencies, intergroup dynamics, and governance responses.

Following Whittemore and Knafl's (2005) methodology, the review proceeded through five stages: problem identification, literature search, data evaluation, analysis, and presentation. The problem was defined as identifying socio-informational patterns of mis/disinformation in Malaysia's

hyper-connected, multiethnic society. Accordingly, the inclusion criteria required that selected works be related to fake news, misinformation, or disinformation in the Malaysian context. Journal articles were located through purposive searches in Google Scholar, JSTOR, ProQuest, and ScienceDirect using Boolean combinations of keywords (e.g., "misinformation" OR "disinformation" OR "fake news" AND "social media" AND "Malaysia"). Grey literature, including news articles and policy reports, was supplemented through targeted Google searches using variations of these terms and filters.

28 sources published between 2015 and 2025 were selected, comprising 16 peer-reviewed journal articles, one book chapter, 10 news articles, and one academic report. Peer-reviewed works formed the analytic core of the review, while grey literature (news articles, report) was incorporated to provide contextual and policy-relevant insights into Malaysia's evolving mis/disinformation landscape. This mixed selection provided sufficient conceptual coverage, emphasizing depth and relevance over exhaustive inclusion, consistent with integrative review guidelines that encourage methodological diversity and the synthesis of diverse insights rather than a systematic inventory of all studies (Snyder, 2019; Torraco, 2005; Whittemore & Knafl, 2005).

Data analysis followed Whittemore and Knafl's (2005) constant comparison method. Extracted findings were organized into matrices capturing variables such as type of mis/disinformation, dissemination platform, socio-political context, and responses. Iterative comparison enabled clustering of recurring codes into higher-order categories, which were then refined and verified against primary sources. Divergent findings were critically examined to preserve complexity rather than being excluded. This process yielded three overarching socio-informational patterns, which will be advanced further in the following section.

The theoretical perspectives outlined earlier directly shaped this analysis: diffusion and confirmation bias informed the identification of technological-cognitive amplification mechanisms; Intergroup Threat Theory guided attention to ethnic and religious framings; and Securitization Theory provided a lens for interpreting state narratives and regulatory responses. This combination of an integrative review methodology and theory-driven thematic analysis ensured both conceptual richness and methodological rigor, generating an informed understanding of how mis/disinformation circulates and is contested in Malaysia.

RESULTS AND DISCUSSION

The integrative review reveals three core socio-informational patterns shaping mis/disinformation in Malaysia: (1) digital diffusion and cognitive reinforcement; (2) disinformation as a tool of intergroup threat; and (3) securitizing disinformation through law and contestation.

Theme 1: Digital Diffusion and

Cognitive Reinforcement

The integrative review reveals that mis/disinformation circulates rapidly and embeds itself structurally in Malaysia, where algorithmic amplification converges with cognitive tendencies to reinforce falsehoods. Digital platforms, particularly social media and private messaging networks such as WhatsApp, function as echo chambers that continuously normalize belief-consistent narratives rather than hosting fleeting viral content (Akram et al., 2023; Mohd Yatid, 2019; Rozwan et al., 2023). This technological-psychological interplay is further magnified by Malaysia's multilingual media ecosystem, which is characterized by low institutional trust (Nain, 2023; Radue, 2019). In such an environment, skepticism toward mainstream news outlets drives citizens to rely on peer-to-peer networks that lack editorial oversight, enabling mis/disinformation to traverse Malay, Chinese, and Tamil communities with minimal resistance. The combination of linguistic segmentation and weak institutional credibility provides fertile ground for narratives to be tailored to distinct communities and reinforced within self-contained echo chambers.

Several studies identified in the integrative review (Daud & Zulhuda, 2020; Mohd Yatid, 2019; Zaifuddin et al., 2023) identify these digital platforms as an ideal ground for disseminating false content, owing to their accessibility, low cost, and anonymity. For example, Daud and Zulhuda (2020) highlight that social media's appeal for spreading false content lies in its affordability, anonymity, and the way virality lends perceived credibility, particularly when public figures are involved. Affordability and perceived credibility, therefore, accelerate the spread of false narratives, embedding them in everyday discourse. The Malaysian Communications and Multimedia Commission, the nation's regulatory body overseeing telecommunications, multimedia, online content, and digital industry compliance, recorded over 1,000 instances of false news from 2013 to 2017 (Bernama, 2017, as cited in Daud & Zulhuda, 2020). Moreover, more recent official figures indicate a continued and intensified trend, with 3,285 complaints related to fake news received between 2020 and May 2022 (Ministry of Communications and Digital, 2022). COVID-19 further amplified these dynamics, with widespread falsehoods concerning vaccines and health measures proliferating across Facebook and WhatsApp (Zaifuddin et al., 2023). These incidents show how mis/disinformation exploits moments of collective uncertainty, such as elections or pandemics, where demand for rapid information outpaces institutional capacity to supply accurate guidance.

To interpret these patterns, Rogers' (1983) Diffusion of Innovations Theory offers a compelling lens. Like any social innovation, mis/disinformation spreads when perceived as advantageous, compatible with existing beliefs, and easy to adopt. Mis/disinformation aligns well with these characteristics on platforms where emotionally charged or sensationalist content attracts more engagement. Moreover, digital platforms reward such content through algorithmic amplification, transforming virality into a proxy for credibility. Mohd Yatid (2019) further highlights how social

validation mechanisms, such as likes, retweets, and shares, act as perceived endorsements, motivating users to spread content with little critical scrutiny. When these mechanisms are exploited through bots and automation, the circulation of falsehoods escalates from incidental sharing to a coordinated ecosystem. This indicates that disinformation in Malaysia is not only opportunistic but increasingly systematic, driven by actors who strategically leverage digital infrastructures for political, economic, or ideological ends.

Complementing this diffusion model is confirmation bias, whereby users preferentially accept and share content that aligns with their preexisting beliefs, a tendency first demonstrated by Wason (1960). Building on this foundation, Meserole (2018) describes social media as a "confirmation bias machine," where algorithms reinforce users' existing beliefs by curating like-minded content, fostering echo chambers. In open networks like Facebook and X, users navigate environments where speed and visibility are prioritized over accuracy as both consumers and amplifiers of content. The interaction between human cognitive tendencies and platform algorithms creates a feedback loop in which falsehoods are spread, legitimized, and normalized. Consequently, as users engage more with content that affirms their beliefs, they unintentionally increase the reach and credibility of false narratives.

In the Malaysian context, this phenomenon is exacerbated by certain socio-cultural and institutional factors. With a digital penetration rate of 97.7%—covering almost the entire Malaysian population (Kemp, 2025) and a fragmented linguistic and media ecosystem (Lee & Wong, 2020), the country faces significant challenges in implementing effective regulations. Moreover, trust in mainstream media remains low (Nain, 2023; Radue, 2019), often pushing citizens to rely on peer-shared information, particularly within private networks like WhatsApp. Despite being personal and reliable, these channels lack editorial oversight, allowing mis/disinformation to easily spread undetected. As Daud and Zulhuda (2020) observe, enforcement gaps and evolving digital tactics have made it difficult for legal and administrative mechanisms to keep up. This is further complicated by Malaysia's multilingual environment, where mis/disinformation can spread in various vernacular languages, including Bahasa Malaysia, Mandarin, Tamil, and English, often evading detection and response efforts. Here, the interplay between weak institutional trust and linguistic plurality illustrates why interventions focusing only on technological solutions are insufficient without parallel media literacy and trust-building efforts.

The literature depicts an information ecosystem under strain, where the escalation of false content is not just a technological problem, but a socio-behavioral and institutional one. Cognitive biases, platform logics, linguistic fragmentation, and weak institutional trust collectively enable the amplification and normalization of mis/disinformation, suggesting that fact-checking alone is insufficient. It demands structural interventions in regulation, platform accountability, and media and information literacy (MIL). In this sense, Malaysia exemplifies how mis/disinformation should not be conceptualized as isolated falsehoods but as a durable socio-informational process embedded within everyday communication structures.

Theme 2. Disinformation as a Tool of Intergroup Threat

A body of research highlights how disinformation campaigns in Malaysia strategically weaponize ethnic and religious narratives to amplify intergroup anxieties and mobilize hostility. This theme reflects the social layer of the framework, showing how Malaysia's long-standing Malay-Muslim-centric political dominance fosters social stratification and fragments the media environment into polarized echo chambers (Akram et al., 2023; Radue, 2019). These conditions facilitate negative political campaigning online and the rise of polarized communities that can escalate into offline violence. In this context, disinformation emerges not as incidental noise but as a deliberate political instrument exploiting Malaysia's ethnicized social order (Akram et al., 2023; Wan Hashridz et al., 2025; Radue, 2019).

Stephan and Stephan's (2000) Intergroup Threat Theory offers a practical explanatory framework to better understand the psychological mechanisms underpinning these divisions. The theory posits that perceived realistic threats to resources, safety, or political power, as well as symbolic threats to cultural values and identity, increase anxiety and hostility between groups (Stephan & Stephan, 2000). In Malaysia, disinformation campaigns amplify these perceived threats by portraying out-groups and political opponents as existential dangers to the dominant Malay-Muslim identity and socio-political order (Akram et al., 2023; Radue, 2019). This process not only sustains majority-group narratives but also conditions minority communities to navigate a media landscape marked by suspicion and exclusion. The rise in intergroup threat perceptions fosters defensive behaviors and exclusionary attitudes, deepening polarization and entrenching segregated media consumption patterns (Radue, 2019).

What is particularly critical here is that these media practices normalize mistrust, turning suspicion of those perceived as outsiders into an everyday reflex rather than an occasional response. This everyday suspicion helps explain why disinformation often "sticks" even after being corrected, because it taps into existing social hierarchies and deep-seated feelings of vulnerability. In this sense, the persistence of falsehoods resonates with the dynamics outlined in Theme 1, where cognitive biases reinforce preferred narratives; here, however, the reinforcement is not only psychological but also socially structured through intergroup tensions. When amplified at scale, this normalization of mistrust is not limited to online discourse but spills into offline conflict. A striking example is Malaysia's 2015 Plaza Low Yat riot, where false social media posts fueled racial tensions and unrest, prompting police warnings against misinformation to safeguard public harmony ("Social media posts on unrest all false," 2015).

The above analysis aligns with public perceptions and emerging technological challenges, as highlighted in Neo (2021) and Neyazi et al. (2024). Neo (2021) identifies mis/disinformation as a significant threat to Malaysia's social harmony, with many respondents viewing fake news as a greater risk than national security, the economy, or the political system. This illustrates that disinformation is socially consequential not only because of its falsity but also because it resonates with pre-existing identity fault lines, amplifying fears that already structure Malaysian politics. Such campaigns intensify intergroup anxiety and accelerate the spread of false content at scale, further deepening polarization within Malaysia's already fragile and deeply divided socio-political landscape (Zannettou et al., 2018).

Besides, emerging AI-driven forms of disinformation, such as deepfakes, present new challenges by intensifying socio-political polarization (Neyazi et al., 2024). AI-generated deepfakes are making it increasingly challenging to distinguish between authentic and fabricated information (Vaccari & Chadwick, 2020). The increasing accessibility of AI tools and synthetic media technologies has fueled a growing trend in their use for disinformation, with malicious actors exploiting these tools to create convincing yet misleading content on a large scale (Chesney & Citron, 2019; West, 2021). Neyazi et al. (2024) highlight how these synthetic media are weaponized in Malaysia to spread online hate against opposition politicians and minority groups, thereby exacerbating existing divisions in a society historically marked by racially charged politics (Weiss, 2013). Such false information, as Sabri and Abdul Rahim (2023) caution, carries significant risks of social discord and violence, constituting what Islamic ethics identifies as *fasad*, a form of mischief or societal corruption that disrupts social harmony, especially in multiethnic contexts, thereby underscoring the urgent need for proactive and inclusive countermeasures.

In light of these developments, there is an urgent need for strategies that strengthen intergroup understanding and build digital resilience to reduce the social impact of disinformation. This requires not only technical interventions but also efforts to foster empathy and recognition across ethnic lines, so that communities become less susceptible to narratives that frame "the other" as a threat. At the same time, governments often turn to legal responses, which will be examined in the following theme on the securitization of disinformation in Malaysia.

Theme 3: Securitizing Disinformation Through Law and Contestation

The Malaysian government's evolving legal and institutional measures against fake news reveal how mis/disinformation has been securitized as an existential threat that justifies extraordinary regulatory action. In this context, mis/disinformation becomes a site of political contestation, where the state and societal actors vie to define truth and legitimacy. Laws such as the Communications and Multimedia Act 1998 and the Printing Presses and Publications Act 1984, alongside pandemic-era emergency ordinances targeting COVID-19 misinformation illustrate securitization moves that seek to legitimize state control under the banner of national stability. These measures show how communicative power and political authority intersect, revealing that narrative control is central to governance in a highly mediated society. This theme reflects the political-institutional layer of the framework, where disinformation is not only a communicative issue but also a securitized one, contested between state actors seeking legitimacy and civil society pushing back against excessive state control.

A good case of such securitization was the Anti-Fake News Act 2018 (AFN), one of the earliest stand-alone laws in the world to criminalize fake news (Zainudin et al., 2024). Enacted by then-Prime Minister Najib Razak amid the 1Malaysia Development Berhad (1MDB) scandal, its introduction immediately raised concerns about its political motivations (Human Rights Watch, 2018; "Malaysia accused of muzzling critics," 2018). While the government attempted to justify AFN as a necessary tool to protect public order and national security, the law encountered strong opposition from civil society, media practitioners, and opposition parties, who viewed it as overly broad and a means to silence dissent (Zainudin et al., 2024; Radue, 2019). The contested nature of this intervention was exemplified by AFN's swift repeal following the 2018 general election, coinciding with a historic political transition marked by the opposition's first-ever rise to federal power in Malaysia (Lee, 2025).

This dynamic can be meaningfully understood through the lens of Securitization Theory (Wæver, 1995), which explains how authorities frame particular issues, such as fake news, as threats that justify extraordinary legal and political interventions. In Malaysia's case, the AFN was presented by the government as necessary for national security and public order, casting misinformation as an existential threat to societal stability. However, securitization is contingent on audience acceptance; laws and official narratives alone cannot render an issue a legitimate security concern. Civil society groups, journalists, and opposition parties actively challenged the framing of fake news as a national threat, instead portraying the AFN as politically motivated censorship linked to the 1MDB scandal. These actors undermined the law's legitimacy by mobilizing public opinion and leveraging legal and media discourse, demonstrating how securitization is a negotiated process rather than a unilateral imposition. The post-election repeal of the AFN exemplifies this dynamic: without broad societal acceptance, even legally enforceable measures can fail to institutionalize a security narrative. This case highlights that securitizing moves in pluralistic and politically competitive contexts are particularly fragile, as multiple actors can contest, reinterpret, or neutralize state claims of existential threat.

The Malaysian experience with legal frameworks like the AFN unfolds within a broader environment shaped by longstanding skepticism toward state regulation of speech and media. This skepticism is rooted in the nation's history of media control and political suppression, exemplified by draconian laws such as the Internal Security Act 1960, which instilled enduring public apprehension toward government-

led information governance (Wan et al., 2023). Such historical memory amplifies contemporary contestations, as citizens are inclined to view securitizing claims with suspicion, especially when framed in politically charged contexts. Neo (2021) further observes that discourses on fake news are often presented as threats to social harmony in ways that deliberately tap into Malaysia's sensitive ethnic and religious fault lines. Consequently, interventions like the AFN are frequently perceived not as neutral safeguards of public order but as politically motivated tools to maintain the ruling party's dominance. This historical and socio-political backdrop highlights the fragility of securitization efforts in Malaysia, demonstrating that state attempts to assert authority over information are always mediated by public memory, societal mistrust, and the capacity of civil society and media actors to contest legal and political claims.

Recognizing that purely legislative responses to mis/disinformation are insufficient in Malaysia, scholarly discourse has emphasized the need for complementary, citizen-centric approaches. Saharrudin et al. (2023) stress the necessity to equip the public with competencies to critically assess and verify information, while Sabri and Abdul Rahim (2023) call for investment in MIL, practical training, and access to verification tools. In a comparative study, Neyazi et al. (2024) highlight how Singapore has prioritized digital literacy and public education on risks posed by emerging technologies such as AI and deepfakes. This perspective highlights that effective counter-disinformation strategies require institutional legitimacy and citizen engagement.

During the COVID-19 pandemic, Malaysia's Ministry of Health demonstrated how trust-based, communicative strategies could complement or replace punitive legal measures by using social media to share daily statistics, prevention videos, and official updates, while actively debunking false news (Zaifuddin et al., 2023). This approach underscores that legitimacy can be cultivated through transparency and responsiveness rather than coercion alone, highlighting the value of citizen-centric engagement in countering mis/disinformation. To ensure effectiveness in the long term, these strategies should be institutionalized through sustained educational programs, MIL initiatives, and continuous public engagement, enabling Malaysia to navigate the evolving landscape of misinformation.

CONCLUSION

The integrative review reveals that mis/disinformation in Malaysia operates as a complex, multi-layered socio-informational phenomenon, shaped simultaneously by technological, cognitive, social, and political dynamics. Beyond the mere circulation of falsehoods, mis/disinformation functions as a systemic mechanism that reinforces cognitive biases, exploits intergroup anxieties, and engages contested state-society power dynamics. By integrating digital diffusion, intergroup threat, and securitization perspectives, the study highlights how the Malaysian context, characterized by its multilingualism, low institutional trust, and ethnoreligious sensitivities, creates a fertile ground for mis/disinformation and significant barriers to effective regulation. The findings emphasize that countering mis/disinformation cannot rely solely on technological or legal interventions but requires a holistic approach considering psychological, social, and institutional dimensions. The study provides a complex lens through which policymakers, media practitioners, and scholars can understand the persistent vulnerabilities within Malaysia's hyper-connected, multiethnic society and design interventions that balance state oversight with citizen engagement, MIL, and social cohesion.

While offering a rich synthesis of theoretical and empirical insights, this study presents several limitations. Drawing on 28 purposively selected English-language sources, it potentially excludes non-English scholarship and additional unpublished materials. Its qualitative, interpretive nature limits statistical generalizability, and reliance on secondary sources introduces potential biases and methodological heterogeneity (Whittemore & Knafl, 2005). Additionally, the chosen theoretical frameworks may highlight specific dynamics while underrepresenting others, leaving space for alternative perspectives in future research.

Future studies could focus on socio-informational patterns identified in this review by employing empirical methods such as surveys, interviews, or digital ethnography to validate and extend the findings. Longitudinal research could track the evolution of mis/disinformation, particularly AI-driven content and deepfakes. Moreover, comparative studies across Southeast Asia would help identify region-specific drivers and mitigation strategies. Finally, evaluating citizen-centric interventions, such as MIL programs and community-based counter-disinformation initiatives, could generate actionable insights for policy and practice in Malaysia's hyper-connected, multiethnic context.

REFERENCES

Akram, M., Nasar, A., & Arshad-Ayaz, A. (2023). A systematic review for netizens' response to the truth manipulation on social media. *Knowledge Management & E-Learning: An International Journal, 15*(2), 322–342. https://doi.org/10.34105/j.kmel.2023.15.018

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives, 31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. New York: Anchor Books.

Briggs, A., & Burke, P. (2018). *A social history of the media: From Gutenberg to the Internet* (4th ed.). Polity Press.

Campbell, R., Martin, C. R., & Fabos, B. (2020). *Media & culture: Mass communication in a digital age* (12th ed.). Bedford/St. Martin's.

Chesney, B., & Citron, D. (2019). Deepfakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1820. https://doi.org/10.15779/Z38RV0D15J

Daud, M., & Zulhuda, S. (2020). Regulating the spread of false content online in Malaysia: Issues, challenges and the way forward. *International Journal of Business and Society, 21*(S1), 32–48. https://www.ijbs.unimas.my/images/repository/pdf/Vol21-S1-paper3.pdf

Eisenstein, E. L. (2012). *The printing revolution in early modern Europe* (2nd ed.). Cambridge University Press.

Fake news, online provocations dropped after GE15 campaign period. (2022, November 24). *Bernama*. https://www.bernama.com/en/news.php?id=2141842

Fallis, D. (2015). What is disinformation? *Library Trends, 63*(3), 401–426. https://doi.org/10.1353/lib.2015.0014

GE15: EC tells voters to ignore fake news about polling day dress code banning slippers, shorts. (2022, November 19). *Malay Mail*. https://www.malaymail.com/news/malaysia/2022/11/19/ge15-ec-tells-voters-to-ignore-fake-news-about-polling-day-dress-code-banning-slippers-shorts/40648

George, J. (2022). *Fortifying the ballot: Addressing disinformation and electoral integrity in Malaysia* (ISEAS-Yusof Ishak Institute Perspective No. 72). ISEAS-Yusof Ishak Institute. https://www.iseas.edu.sg/wp-content/uploads/2024/08/ISEAS_Perspective_2024_72.pdf

He, D., & Sannusi, S. N. (2024). Motives of sharing fake news and effects on mental health of social media users: A meta-analysis. *Jurnal Komunikasi: Malaysian Journal of Communication, 40*(2), 188–204. https://doi.org/10.17576/JKMJC-2024-4002-11

Human Rights Watch. (2018, March 30). *Malaysia: Drop proposed 'fake news' law.* https://www.hrw.org/news/2018/03/30/malaysia-drop-proposed-fake-news-law

Karas, Z. (2024). Effects of AI-generated misinformation and disinformation on the economy. *Journal of Science & Technology, 12*, 2349–2360. https://dergipark.org.tr/en/download/article-file/4162778

Kemp, S. (2025, March 3). Digital 2025: Malaysia. *DataReportal*. https://datareportal.com/reports/digital-2025-malaysia

Krznaric, R. (2024). *Fake news was invented 500 years ago – We must learn from its history. openDemocracy.* https://www.opendemocracy.net/en/fake-news-invention-witch-hunt-history-for-tomorrow-book-roman-krznaric

Lee, K. T. W. (2025). Malaysia's media freedom: A new dawn under Anwar Ibrahim or history repeating? In P. V. Kien, D. A. Duc, N. T. T. Hang, & N. K. Nga (Eds.), *Digital convergence in media: Vietnam and transnational perspectives* (pp. 445–470). Nomos. https://doi.org/10.5771/9783748954286-445

Lee, K. T. W., & Wong, W. W. (2020). Framing Jawi-Khat move: A comparative analysis of Chinese, English and Malay-language newspapers in Malaysia. *Malaysian Journal of Communication, 36*(4), 194–210. https://doi.org/10.17576/JKMJC-2020-3604-12

Levack, B. P. (2016). *The witch-hunt in early modern Europe* (4th ed.). Routledge.

Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour, 5*(3), 337–348. https://doi.org/10.1038/s41562-021-01056-1

Lund, K. (2022, November 23). PDRM confirms that viral TikTok photos of a "riot" in Perak regarding GE15 are fake. *Juice*. https://juiceonline.com/tiktok-ge15-riot-perak-fake

Malaysia accused of muzzling critics with jail term for fake news. (2018, March 26). *The Guardian*. https://www.theguardian.com/world/2018/mar/26/malaysia-accused-of-muzzling-critics-with-jail-term-for-fake-news

Meserole, C. (2018, May 9). How misinformation spreads on social media—and what to do about it. *The Brookings Institution*. https://www.brookings.edu/articles/how-misinformation-spreads-on-social-media-and-what-to-do-about-it

Ministry of Communications and Digital. (2022). *Fake news spreaders deserve heavier penalty*. https://www.komunikasi.gov.my/awam/berita/23980-fake-news-spreaders-deserve-heavier-penalty

MOH acted on 14 viral fake news posts for Covid misinformation. (2022, March 3). *CodeBlue*. https://codeblue.galencentre.org/2022/03/moh-acted-on-14-viral-fake-news-posts-for-covid-misinformation

Mohd Yatid, M. (2019). Truth tampering through social media: Malaysia's approach in fighting disinformation & misinformation. *IKAT: The Indonesian Journal of Southeast Asian Studies, 2*(2), 203–230. https://doi.org/10.22146/ikat.v2i2.40482

Nain, Z. (2023, June 14). *Malaysia: Three years of the pandemic, lockdowns, and political and economic uncertainty deepen worries for print media*. Reuters Institute Digital News Report. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/malaysia

Neo, R. (2021). The failed construction of fake news as a security threat in Malaysia. *Contemporary Politics, 27*(3), 316–335. https://doi.org/10.1080/13569775.2021.1884397

Neyazi, T. A., Nadaf, A. H., Tan, K. E., & Schroeder, R. (2024). Does trust in government moderate the perception towards deepfakes? Comparative perspectives from Asia on the risks of AI and misinformation for democracy. *Government Information Quarterly, 41*(2), 101980. https://doi.org/10.1016/j.giq.2024.101980

Radue, M. (2019). Harmful disinformation in Southeast Asia: "Negative campaigning", "information operations" and "racist propaganda" – Three forms of manipulative political communication in Malaysia, Myanmar, and Thailand. *Journal of Contemporary Eastern Asia, 18*(2), 68–89. https://doi.org/10.17477/jcea.2019.18.2.068

Rafizi's Tabung Haji claims example of why we need fake news law. (2018, April 3). *New Straits Times.* https://www.nst.com.my/news/nation/2018/04/3520 17/rafizis-tabung-haji-claims-example-why-we-need-fake-news-law

Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed.). New York, NY: Free Press.

Ross, A. S., & Rivers, D. J. (2022). Discursive deflection: Accusation of "fake news" and the spread of mis- and disinformation in the public sphere. *Social Media + Society, 8*(2), 1–12. https://doi.org/10.1177/2056305118776010

Rozwan, I. D., Husin, A. K., Nazri, A. S., Zainal, N. A., & Hassan, M. S. (2023). The impact of fake news on political leaders: Reputation, policies and public perception. *e-Journal of Media & Society, 6*(2), 43–54.

Sabri, S. A., & Abdul Rahim, N. R. (2023). Ethical implications of fake news: A meta-analysis through the lens of Islamic teachings. *Al-i'lam: Journal of Contemporary Islamic Communication and Media, 3*(2), 157–168.

Saharrudin, N. F., Ahmad Ghazali, A. H., Ismail, I. A., Arshad, M. M., Mohd Zawawi, J. W., & Wan Mokhtar, W. I. (2023). Reasons for fake news dissemination: A systematic review. *International Journal of Academic Research in Business and Social Sciences, 13*(17), 284–302. https://doi.org/10.6007/ijarbss/v13-i17/19836

Shoib, M. R., Wang, Z., Ahvanooey, M. T., & Zhao, J. (2023). Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models. *AI & Society, 38*(1), 1–7. https://doi.org/10.48550/arXiv.2311.17394

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Social media posts on unrest all false, say police. (2015, July 16). *Malaysiakini.* https://www.malaysiakini.com/news/305323

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–45). Lawrence Erlbaum Associates Publishers.

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining 'fake news': A typology of scholarly definitions. *Digital Journalism, 6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tapsell, R. (2023). Social media and Malaysia's 2022 election: The growth and impact of video campaigning. *Pacific Affairs, 96*(2), 303–322. https://doi.org/10.5509/2023962303

The threat of deepfake to Malaysia. (2019, July 19). *The Malaysian Reserve.* https://themalaysianreserve.com/2019/07/19/the-threat-of-deepfake-to-malaysia

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review, 4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society, 6*(1), 1–13. https://doi.org/10.1177/2056305120903408

Wæver, O. (1995). Securitization and desecuritization. In R. D. Lipschutz (Ed.), *On security* (pp. 46–86). Columbia University Press.

Wan Hashridz Rizal Wan Abu Bakar, Abdul Rauf Ridzuan, Mohd Sufiean Hassan, & Fahmi Zaidi Abd Razak. (2025). Exploring fake news sharing among Malaysian university students: Role of uses and gratification theory. *Environment-Behaviour Proceedings Journal, 10*(SI26), 21–26. https://doi.org/10.21834/e-bpj.v10iSI26.6788

Wan, C., Waheed, M., Zawawi, J. W. M., & Hellmueller, L. (2023). The framing of anti-fake news law in Malaysian newspapers. *International Journal of Academic Research in Business and Social Sciences, 13*(3), 1208–1224. http://dx.doi.org/10.6007/IJARBSS/v13-i3/16557

Wardle, C. (2017, February 16). Fake news. It's complicated. *First Draft News.* https://firstdraftnews.org/articles/fake-news-complicated

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. *Council of Europe.* https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c018.1502047

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*(3), 129–140. https://doi.org/10.1080/17470216008416717

Weiss, M. L. (2013). *Protest and possibilities: Civil society and coalitions for political change in Malaysia*. Stanford University Press.

West, D. M. (2021). *How to combat fake news and disinformation*. Brookings Institution. https://www.brookings.edu/research/how-to-combat-fake-news-and-disinformation

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing, 52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

World Health Organization. (2020, February 2). *Novel coronavirus (2019-nCoV): Situation report - 13*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200202-sitrep-13-ncov-v3.pdf

Zaifuddin, N. S., Adzharuddin, N. A., Osman, M. N., & Mohd Zawawi, J. W. (2023). Fake news during COVID-19 in Malaysia: How Ministry of Health encountered fake news contagion in Facebook context. *International Journal of Academic Research in Business and Social Sciences, 13*(12), 5833–5843. https://doi.org/10.6007/ijarbss/v13-i12/20392

Zainudin, J., Ali, N. M., Smeaton, A. F., & Ijab, M. T. (2024). Intervention strategies for misinformation sharing on social media: A bibliometric analysis. *IEEE Access, 12*, 140359–140379. https://doi.org/10.1109/ACCESS.2024.3469248

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2018). The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality, 11*(3), 1–37. https://doi.org/10.1145/3309699

ORCID

Kenneth Tze Wui Lee https://orcid.org/0000-0002-9959-4640

ABOUT THE AUTHORS

Kenneth Tze Wui Lee is a Senior Lecturer and Head of the Department of Mass Communication at the Faculty of Creative Industries, Universiti Tunku Abdul Rahman, Malaysia. His research centers on media, digital communication, identity construction, and nation-building, particularly in Malaysia and Southeast Asia.

Zhi Xin Sau is pursuing her undergraduate studies at the Faculty of Creative Industries, Universiti Tunku Abdul Rahman, Malaysia. Her academic interests revolve around social media and digital culture. This paper reflects a collaborative effort beyond her formal coursework.