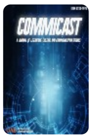
UAD
Universitas
Ahmad Dahlan

COMMICAST
Vol. 6, No. 2, September 2025, pp. 263-279
DOI: https://doi.org/10.12928/commicast.v6i2.13673
ISSN: 2723-7672



Lessons from the social Dilemma: BSSN's social cybersecurity strategy addressing information disorders

[1] Abid Prayoga Hutomo*, [2] Andre Noevi Rahmanto, [3] Sudarmo

Faculty of Social and Political Sciences, Universitas Sebelas Maret, Surakarta, 57126, Indonesia
[1] abid@student.uns.ac.id*; [2] andre@staff.uns.ac.id; [3] sudarmo@staff.uns.ac.id
*Correspondent email author: abid@student.uns.ac.id

ARTICLE INFO

Article history
Received 2025-06-17
Revised 2025-07-30
Accepted 2025-08-26

Keywords
Communication strategy
Disorder of information
Disinformation
Social cybersecurity
Social dilemma

ABSTRACT

The proliferation of misinformation, disinformation, and Malinformation on social media poses a serious threat to public trust and national security. Social media algorithms, as illustrated in The Social Dilemma, inadvertently amplify false and sensational content, fostering polarization and societal vulnerability. This study aims to analyze the strategy of Indonesia's National Cyber and Crypto Agency (*Badan Siber dan Sandi Negara*/BSSN) in addressing information disorder through a participatory digital literacy approach. Its main contribution lies in providing both academic insights and policy recommendations for an ethical, adaptive, and evidence-based model of social cybersecurity governance. This research employed a descriptive qualitative method, combining documentary analysis of The Social Dilemma, a review of official BSSN documents, and an in-depth interview with a BSSN official. The data were processed using thematic coding and triangulated across multiple sources to ensure credibility and validity. The findings reveal that BSSN implements the EMILIE framework, Encouragement, Measurement, Involvement, Literacy, and Empowerment, which strengthens digital literacy, promotes stakeholder engagement, and develops ethical monitoring systems while safeguarding civil rights. This framework has proven effective in raising public awareness and resilience against disinformation, although challenges remain, such as the rapid spread of harmful content, reliance on platform cooperation, and limited institutional resources. In conclusion, participatory and literacy-based approaches to social cybersecurity are essential in countering digital information disorder. BSSN's strategy demonstrates that fostering multi-stakeholder collaboration and community empowerment can mitigate cyber threats while ensuring ethical and legal protection of citizens.

1. Introduction

Social cybersecurity has emerged as a strategic subdomain within the national security framework, significantly shaping future conventional and unconventional warfare. As a relatively new discipline, social cybersecurity aims to understand, anticipate, and manage human behavior and the social, cultural, and political dynamics that are increasingly mediated by cyberspace (Carley, 2020). The primary focus of this field is the development of social cyber infrastructure that can safeguard fundamental social values in an information landscape riddled with real and potential social threats (Wardle & Derakhshan, 2017).

In Indonesia, the cybersecurity threat landscape has evolved from conventional technical attacks such as hacking and malware dissemination to social engineering assaults that exploit disinformation, digital propaganda, and online manipulation of public opinion (Awadallah et al., 2024; Chaudhuri et al., 2025). This shift is propelled by two fundamental changes in communication: the elimination of geographical boundaries through digital technology, which facilitates external intervention without physical presence, and the decentralization of information flow, thereby lowering dissemination barriers and complicating attribution to malicious actors. As depicted in Fig. 1, information disorder within the digital domain presents in various forms: misinformation, such as fake connections or misleading content; disinformation, including manipulated, fabricated, or deceptive material; and information, which comprises fraud, cyberbullying, leaks, hate speech, or the non-consensual distribution of intimate images. Misinformation involves inaccurate information that the speaker believes to be true, while disinformation consists of false information intentionally fabricated to deceive or manipulate public opinion (Roshanaei et al., 2024). False information, on the other hand, may contain elements of truth but is deliberately presented to cause harm. As a result, such information disruptions are now spreading rapidly due to high connectivity, low barriers to publication, and source anonymity, presenting unique challenges for authorities seeking effective and timely responses (Wardle & Derakhshan, 2017).

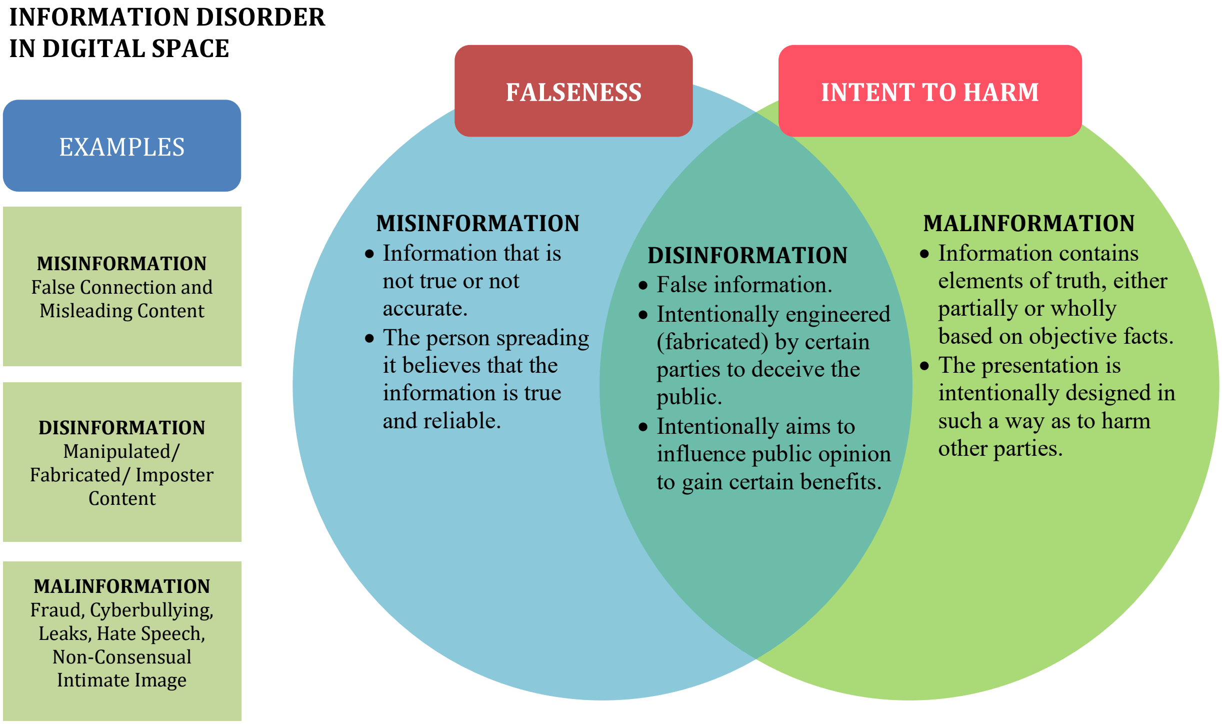

INFORMATION DISORDER
IN DIGITAL SPACE
EXAMPLES
MISINFORMATION
False Connection and Misleading Content
DISINFORMATION
Manipulated/ Fabricated/ Imposter Content
MALINFORMATION
Fraud, Cyberbullying, Leaks, Hate Speech, Non-Consensual Intimate Image
FALSENESS
INTENT TO HARM
MISINFORMATION
• Information that is not true or not accurate.
• The person spreading it believes that the information is true and reliable.
DISINFORMATION
• False information.
• Intentionally engineered (fabricated) by certain parties to deceive the public.
• Intentionally aims to influence public opinion to gain certain benefits.
MALINFORMATION
• Information contains elements of truth, either partially or wholly based on objective facts.
• The presentation is intentionally designed in such a way as to harm other parties.

Fig. 1. Information Disorder in Digital Space

Source: Wardle C. Essential Guide to Understanding Information Disorder

The urgency of this challenge is clearly illustrated by The Social Dilemma McDavid, (2020) , which demonstrates how social media algorithms are engineered to exploit human psychology, thereby maximizing engagement while inadvertently amplifying misinformation and social

https://doi.org/10.12928/commicast.v6i2.13673

polarization. The technology experts featured in the film emphasized the paramount importance of implementing comprehensive interventions (Iddianto & Azi, 2022). Empirical cases in Indonesia, such as the widespread circulation of manipulated videos including deepfakes depicting President Joko Widodo speaking fluent Mandarin Kominfo, (2023), highlight the sophisticated nature of current threats. National Cyber and Crypto Agency (BSSN) and the Ministry of Communication and Digital (Komdigi) handled 5,092,559 cases of harmful content on websites and 2,300,327 on social media platforms throughout 2024, with an average of approximately 7,900 negative content items created daily. The most common types are gambling content (3,824,202 cases) and pornography (1,221,243 cases). Meta and X (Twitter) were identified as the primary platforms used for dissemination. This data further emphasizes that modern cyberattacks are now strategically targeting social and psychological domains, undermining the integrity of public information and manipulating collective opinion with broad political and social consequences (Husandani et al., 2025).

Globally recent research has delineated the intricate landscape of information disorders, advocating for an integrated and interdisciplinary framework (Carley, 2020; Wardle & Derakhshan, 2017). Mulahuwaish et al., (2025) emphasizes that social cybersecurity, in contrast to traditional cybersecurity, concentrates on the human and social aspects of cyber threats such as cyberbullying, cybercrime, and online manipulation. Furthermore, it underscores the critical necessity for adaptive detection techniques and policy measures. For example, Arroyo et al., (2021) comprehensively reviewed detection techniques for social cyberattacks, highlighting ongoing challenges in scalability, detection accuracy, and the management of ethical intervention. Abdo et al., (2025) Highlight the proliferation of bots and malicious campaigns that exploit regulatory and technical gaps, which necessitate more adaptive, cross-sectoral solutions. Critical gaps also persist in algorithmic and AI literacy, with differences in digital skills increasing user vulnerability to manipulation (Chung & Wihbey, 2024; Kong et al., 2024). The swift advancement of synthetic and AI-generated media contributes to this risk, as contemporary enemy attack and defense strategies are frequently insufficient for practical, real-world applications (Cartwright et al., 2025; Cloos et al., 2025; Kong et al., 2024).

In the context of the Indonesian government's initiatives, empirical studies have commenced to examine information exchanges and the dynamics of social networks along with responses to national policies (Abdillah et al., 2024; Maddock-Ferrie, 2022). However, (Surjatmodjo et al., 2024) note that most of this research remains fragmented and descriptive, lacking a comprehensive multi-stakeholder framework for building national resilience. The government’s tangible initiatives, such as the EMILIE approach (Encouragement, Measurement, Engagement, Literacy, Empowerment) and the prospective adoption of a pseudo-panoptic model by the BSSN, exemplify an integrated response to information disruption. Nonetheless, comprehensive academic assessments of their practical effectiveness, scalability, and ethical considerations remain limited. Addressing information disorder entails a complex equilibrium: restricting specific types of information may infringe upon freedom of expression, whereas permitting the unregulated dissemination of falsehoods can jeopardize public health and safety.

Accordingly, the government must formulate strategies that safeguard the well-being of its citizens and uphold democratic principles. Given its pivotal role, the BSSN is uniquely positioned to develop and execute this social cybersecurity strategy. Consequently, this research endeavors to bridge the existing gap by critically analyzing the BSSN approach, including the potential application of the pseudo-panopticon model in mitigating information overload in Indonesia. Specifically, this study addresses two primary questions: (Q1) How do social media algorithms facilitate the dissemination of disorder of information? And (Q2) How can BSSN implement an effective social cybersecurity strategy, and to what extent can the pseudo-panopticon approach be employed to address the disorder of information in Indonesia?

2. Method

This study employs a descriptive qualitative methodology to examine how social media algorithms contribute to the disorder of information, encompassing misinformation, disinformation, and malinformation. Additionally, it analyzes BSSN's strategic responses within the social cybersecurity framework. The qualitative approach is deemed suitable due to the complexity of the phenomenon under investigation, as it facilitates an in-depth exploration and interpretation of context-specific strategies in real-world settings.

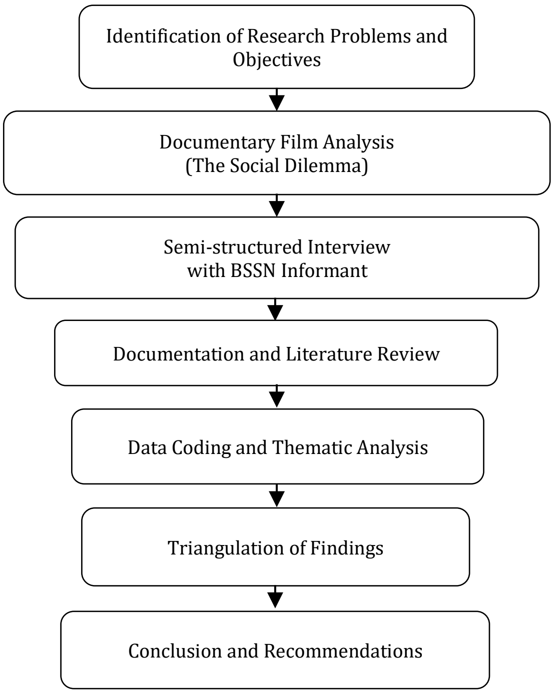

Identification of Research Problems and Objectives
Documentary Film Analysis (The Social Dilemma)
Semi-structured Interview with BSSN Informant
Documentation and Literature Review
Data Coding and Thematic Analysis
Triangulation of Findings
Conclusion and Recommendations

Fig. 2. Research Framework
Source: Processed by researchers

Researchers gathered data using three methods, film content analysis, in-depth interviews, and document review. Firstly, the documentary film The Social Dilemma McDavid, (2020) was analyzed as a critical case study to examine expert insights and algorithmic dynamics that facilitate the dissemination of DOI. Selected quotes and scenes were utilized to underscore the primary arguments articulated by technology developers and researchers. Secondly, semi-structured interviews were conducted on 21 June 2024, with a key informant from BSSN. The key informant, M.A., who serves as the Policy Technical Reviewer in the Directorate of Security Operations and Information Control, Deputy for Cyber Security and Cryptography Operations at BSSN, was selected based on their direct involvement in BSSN's social cyber security program. Although the scope was limited to a single participant due to institutional access constraints, the informant's authoritative position offers credible insights into national cybersecurity operations. Thirdly, internal BSSN documentation, particularly the strategic presentation titled "Social Cyber Security: Realizing Humanistic Social Media in the Political Year 2024 to Maintain National Unity in the Context of

https://doi.org/10.12928/commicast.v6i2.13673

National Security" Kholili, (2025), was reviewed alongside academic literature concerning social cyber security, information disorder, and algorithmic governance (Carley, 2020; Wardle & Derakhshan, 2017).

The data were analyzed using a thematic coding process. The initial phase involved identifying recurrent patterns and pertinent concepts across three data sources. This process was subsequently refined through focused coding into overarching analytical themes related to DOI mechanisms, BSSN policy responses, and the framing of social cybersecurity. To enhance credibility, data triangulation was performed across three sources: documentary analysis, in-depth interviews, and literature review, thereby enabling cross-validation and more robust interpretation of the findings. Additional methodological rigor was maintained through the validation of informants' credibility, systematic documentation of each analytical step, and transparency of the research process.

As seen in Fig.2, to illustrate this approach, the methodological flow was conceptualized as follows: problem identification and research objectives; documentary film analysis, The Social Dilemma; semi-structured interviews with BSSN officials; documentation and literature review; thematic coding and analysis; triangulation of findings; and formulation of conclusions and recommendations. This multi-method design provides a comprehensive framework that supports the research objectives of uncovering the intersection between algorithmic structures and national cybersecurity strategies in the face of contemporary information disruptions.

3. Result and Discussion

Result

This section analyzes the research findings related to questions Q1 and Q2, using data from interviews, the documentary The Social Dilemma, documents, and literature. This approach offers insights into information disorder in digital space and BSSN's strategic responses. It discusses how social media algorithms enable information disorder, BSSN's social cybersecurity strategies, and the potential of the pseudo-panopticon approach in Indonesia's social cybersecurity framework.

The Impact of Social Media Algorithms on the Spread of Disorder of Information

Evidence from the documentary film The Social Dilemma reveals how social media algorithms act as a powerful amplifier of the disorder of information, particularly in the form of misinformation, disinformation, and malinformation. Through insights from former tech industry leaders and digital rights advocates featured in the film, it becomes clear that the design and logic of algorithm-based platforms fundamentally shape how various types of information spread and take root in society.

A central quote from the film summarizes the social threat: 'We tweet, we like, and we share, but what are the consequences of our growing dependence on social media? This documentary-drama hybrid reveals how social media is reprogramming civilization with tech experts sounding alarms about their own creations." (Harris, 2019; Hoehe & Thibaut, 2020). The film further illustrates, through dramatized family scenarios, how individuals, particularly young persons, are exposed to, manipulated by, and even develop addictions to customized content that consolidates existing beliefs, constricts perspectives, and fosters increased distrust in public institutions. Algorithms, regarded as artificial intelligence operating covertly in the background, are accountable not only for disseminating but also for magnifying misinformation-related disturbances (Iddianto & Azi, 2022).

Misinformation pertains to false or misleading information that is disseminated unintentionally, often because individuals perceive it as accurate. Tristan Harris, a former Google Design Ethics

Expert and Co-Founder of the Center for Humane Technology, explains that the 'engagement loop' inherent in algorithms incentivizes content that is sensational or emotionally provocative, regardless of its truthfulness. For instance, some rumors suggest that consuming large quantities of water can eliminate viruses, and there are claims that the government plans and creates the virus, leading to mass actions such as the destruction of cell towers due to rumors asserting that it is not COVID-19 that causes fatalities, but rather the 5G signals emitted. Jeff Seibert, a former Senior Product Director at Twitter, and Justin Rosenstein, Co-Founder of Asana and former engineer at Google and Facebook, further elaborate on how personalized feeds and search results establish 'filter bubbles.' These algorithmic silos reinforce preexisting biases and facilitate the rapid and widespread dissemination of misinformation, which is often believed and propagated by ordinary users.

Disinformation involves creating and spreading false information intentionally to deceive or manipulate public opinion. Tristan Harris states that algorithmic tech has surpassed human weaknesses and created "human checkmate," where trapping people in cycles of addiction, radicalization, and anger. He notes that social media monetizes disinformation, citing COVID-19 rumors spread by engagement-driven algorithms: "We have created tools to disrupt and erode society's structure… We built this and are responsible for changing it," Harris said.

Guillaume Chaslot, CEO of Intuitive AI and a former Google/YouTube engineer, noted that YouTube's recommendation system frequently promoted conspiracy videos such as Flat Earth theories to millions of users, not because of their credibility, but because they generated more clicks and longer viewing times. Similarly, Renee Diresta, Research Manager at the Stanford Internet Observatory, highlighted the "Pizzagate" conspiracy as an example in which Facebook's algorithm actively recommended conspiracy groups to users who had previously expressed interest in related narratives. As emphasized by Sandy Parakilas, a former Facebook Operations Manager, platform models driven by profit not only tolerate but also algorithmically incentivize disinformation, since such content tends to yield higher engagement and advertising revenue.

Malinformation, on the other hand, refers to the use of accurate information presented out of context or with malicious intent, often to damage reputations or incite violence. Cynthia Wong, former Senior Internet Researcher at Human Rights Watch, described how social media platforms in Myanmar were used to disseminate factual details about the Rohingya ethnic group with the explicit aim of inciting hatred and violence, contributing to mass atrocities. Likewise, Roger McNamee, an early Facebook investor and technology critic, underscored the platform's role in U.S. elections, where factual information was selectively reframed and amplified to divide the public or undermine individual reputations. When facts are presented selectively or stripped of context, algorithms amplify their reach, thereby shaping election outcomes and deepening societal polarization.

Insights from BSSN

BSSN offers a robust framework for the comprehension and mitigation of misinformation in Indonesia. According to a comprehensive interview with an official, M.A., dated 21 June 2024, BSSN employs Claire Wardle's typology, which delineates seven categories of false information and disinformation. These classifications serve as guiding principles for BSSN's social media monitoring activities.

> "In defining misinformation and disinformation, BSSN refers to Claire Wardle's journal... There are 7 types, including satire/parody, false connections, misleading content, false context, scam

> content, manipulated content, and fabricated content. BSSN uses these 7 types as a guide in monitoring misinformation and disinformation on social media." (M.A., 21 June 2024).

Particularly in the period preceding the 2024 Indonesian general election, operational and monitoring data from BSSN corroborate the findings outlined in the documentary. For instance, from January to August 2023, BSSN documented 134 instances of false information, 115 instances of disinformation, and 41 instances of misinformation related to political events (Mujib et al., 2023).

Institutional analysis highlights a surge in scams, hate speech, and deepfake content, all exacerbated by the algorithmic architecture of digital platforms. Furthermore, BSSN officials note serious social impacts:

> "The impact of information disorder can cause anxiety, reduce trust levels, affect people's mindsets, and disrupt social stability, which has the potential to trigger social conflicts. One example of disinformation that has been widely discussed this year is a video recording of a conversation between Surya Paloh and Anies Baswedan. This disinformation emerged during Indonesia's presidential election and spread widely through social media, triggering speculation and political commotion. As a result, this causes a division of opinion in society, disrupts the democratic process, and disturbs the atmosphere during the presidential election." (M.A., 21 June 2024).

This research confirms that social media algorithms help spread false information rapidly, as highlighted by The Social Dilemma and expert testimonies. These reinforce information disruption, impacting Indonesia socially with unrest, mistrust, and threats to democracy, notably during elections. BSSN monitoring identifies seven types of disruption and notes rising cases indicating digital ecosystem vulnerabilities.

BSSN's Strategic Measures to Address Disorder of Information

BSSN's role within Indonesia's national cybersecurity ecosystem is both comprehensive and adaptable, demonstrating a definitive commitment not only to regulatory and technical oversight but also to societal resilience. The findings derived from in-depth interviews, when corroborated with official institutional documentation, underscore BSSN's structured and strategic methodology in addressing the evolving threat of information disorder.

The BSSN Mandate, as delineated in "Peraturan Presiden Nomor 47 tahun 2023", designates this agency as the primary entity responsible for coordinating national cybersecurity. This agency bears the responsibility of not only formulating and executing security policies but also ensuring that crisis management, international cooperation, capacity development, and regulatory enforcement are integrated into a cohesive national initiative. As one of the BSSN sources emphasized:

> "BSSN is responsible for coordinating the formulation, implementation, and evaluation of the National Cybersecurity Strategy. This includes developing policies, guidelines, and standards related to cybersecurity in Indonesia. Additionally, BSSN plays a vital role in detecting, identifying, and responding to cyber threats that could potentially endanger critical national information infrastructure. We also prepare for and manage cyber crises through the establishment of comprehensive cyber crisis management procedures, as well as enhancing cybersecurity capacity and capabilities, strengthening international cooperation, developing infrastructure, and enforcing regulations and laws in the cyber field." (M.A, 21 Juni 2024).

This statement underscores BSSN's role as a pivotal entity in fostering cross-sectoral coordination, policy leadership, and operational preparedness. In practice, BSSN addresses information disorder with a layered and adaptive strategy. This is evident in their detection systems, public complaint channels, and sustained literacy campaigns. As confirmed in the interview:

> "The digital media monitoring system is designed to oversee, gather, and manage electronic information from a variety of sources, including online news media and social media platforms. Its primary objective is to identify threats and social cyberattacks circulating within cyberspace. Moreover, BSSN facilitates the reporting of detrimental content by offering recommendations for access termination to the Ministry of Communication and Digital. This initiative aims to extend reach and effectiveness in managing negative content. Additionally, to foster a cybersecurity culture, the organization actively implements educational programs, disseminates valuable information, and conducts security awareness campaigns, particularly concerning the societal threat posed by negative content." (M.A, 21 Juni 2024).

This approach aligns with BSSN's official 2023 presentation Kurniawan, (2023), which outlines essential programs including automated threat monitoring, rapid mechanisms for public reporting, and the EMILIE approach digital literacy framework.

Recognizing the multi-dimensional nature of the disorder of information, BSSN has established strong coordination with both governmental and private sector entities.

> "To identify, detect early, and address disinformation, BSSN facilitates the submission of negative content access termination requests from the public by accepting complaints, providing legal analysis, and submitting access termination requests to social media platforms through coordination with the Ministry of Communication and Digital." (M.A., 21 Juni 2024).

Such coordination is further documented in BSSN's official presentation, which emphasizes the importance of inter-agency collaboration for effective content moderation and enforcement. Cross-sector collaboration involves coordinated efforts with the Ministry of Communication and Digital, law enforcement agencies, and community organizations to report, analyze, and block harmful content. BSSN adopts a comprehensive strategy to enhance the culture of information security, encompassing both technical, social, and educational aspects. One of the key innovations of this policy is the EMILIE approach, which provides strategic direction for all digital literacy initiatives in Indonesia. As emphasized by the key resource person:

> "This program aims to harmonize the community's capabilities and knowledge concerning technological advancements in Indonesia. The strategy to enhance the culture of information security is implemented through the EMILIE approach, which comprises Encouragement: motivating all stakeholders to participate actively; Measurement: evaluating the level of awareness; Involvement: engaging all societal strata; Literacy: improving digital knowledge and skills; and Empowerment: strengthening the capacity of the academic and business sectors." (M.A., 21 June 2024).

The quote emphasizes that EMILIE is not just an acronym, but a framework that encompasses all aspects of participation, measurement, engagement, education, and empowerment to strengthen national cybersecurity resilience. BSSN's documentation further elucidates assessment procedures aimed at remote communities, as well as post-campaign surveys conducted to evaluate changes in digital literacy and security awareness (Kurniawan, 2023). The digital literacy program involves the implementation of the EMILIE approach, which integrates campaigns, webinars, bulletins, and educational materials to foster a security-conscious digital culture.

The implementation of digital literacy initiatives under the coordination of BSSN is conducted systematically, grounded in the specific needs identified in the field. This process commences with

https://doi.org/10.12928/commicast.v6i2.13673

the mapping of these needs and proceeds through impact assessments, thereby ensuring that each educational program remains pertinent and practical. The speaker elucidated:

> "Some great examples of digital literacy activities by BSSN include campaigns, socialization efforts, and sharing cybersecurity bulletins. Before launching these initiatives, BSSN thoughtfully assesses remote communities to find the best ways to deliver learning materials, such as creating content for their social media platforms. After the campaigns or socialization efforts, they also gather feedback through surveys to see how well the communities have understood the digital literacy messages. These steps help ensure everyone stays informed and confident with digital skills." (M.A., 21 June 2024).

This quotation emphasizes that each stage of BSSN's digital literacy initiative, including planning, content refinement, and outcome evaluation, is aimed at promoting behavioral change and enhancing the community's capacity to maintain digital security. In addition, BSSN also incorporates program effectiveness evaluations as an integral part of the entire policy and intervention cycle it undertakes. Before and after implementation, each program always goes through planning, monitoring, and evaluation stages based on data and measurable indicators, as emphasized by M.A. below:

> "In each work program, BSSN diligently performs regular monitoring and evaluation. Monitoring helps us see how things are going on the ground, while evaluation helps us measure how well we're meeting our goals and objectives. We also consider Key Performance Indicators (KPIs), such as managing negative content, response times, cybersecurity awareness, and user satisfaction. Plus, we analyze data from monitoring reports and blocking suggestions to guide our future actions, ensuring continuous improvement." (M.A., 21 June 2024).

The aforementioned quotation emphasizes that BSSN relies not solely on formal reports but also utilizes a diverse range of measurement tools, including Key Performance Indicators (KPIs), statistical analyses of content monitoring, and direct surveys of service recipients and the general public. An explicit example of this methodology is the distribution of Service Satisfaction Surveys to relevant stakeholders and the administration of Cybersecurity Awareness Surveys, with the results serving as a foundation for the analysis of public behavior in Indonesia.

In their endeavors to combat the proliferation of disinformation on social media platforms, BSSN encounters several fundamental challenges that impede the efficacy of its interventions. Field findings suggest that the rapidity and volume of negative content generation frequently surpass the capacity of institutions to respond and monitor effectively, leading to a considerable gap in risk management mitigation.

> "Creating an account and sharing negative content is quick, taking only about 10 minutes. However, dealing with negative content can take much longer, anywhere from 4 hours to a full day. Plus, once content is out there, it can easily be shared and uploaded by others, making it challenging to remove disinformation once it's gone viral." (M.A., 21 June 2024).

In addition to temporal challenges, BSSN also encounters limitations in human resources, where the number of monitoring personnel is not proportional to the volume of content produced each day. The risk of human error becomes even greater, especially amid workload pressures and the demands of real-time operations monitoring. "Limited resources can sometimes lead to human errors in supervision. Plus, there are many instances where AI-created or manipulated content might deceive readers if it's not carefully checked," M.A. stated. Furthermore, the emergence of manipulated or AI-produced content can also contribute to the complexity of the issue, as articulated by M.A. as follows: “Many contents are still created or manipulated using AI, and they can mislead readers if the authenticity isn’t carefully checked.”

The official presentation document from BSSN further emphasizes the challenge that disinformation dissemination now transpires across multiple social media platforms and messaging applications, thereby complicating efforts to trace and manage such content. Furthermore, concerning the management of reports and recommendations for the removal of adverse content, the majority still predominantly depend on public complaints and manual validation by BSSN's internal team, which consequently hampers the efficiency of the follow-up actions process.

BSSN’s initiatives in constructing social cybersecurity are not solely grounded in technical considerations but also authentically incorporate social dimensions into each operational practice. The synergy of these two elements establishes the foundation for addressing the complex, multidimensional nature of modern cyber threats. As emphasized by M.A.:

> “BSSN thoughtfully combines social and technical elements in practicing social cybersecurity. For example, when conducting information gathering activities, technical skills are essential for searching for information. At the same time, social aspects play a valuable role in analyzing that information, creating a well-rounded approach obtained.” (M.A., 21 June 2024).

The quotation emphasizes that the effectiveness of BSSN's security strategy lies in its ability to combine cutting-edge data collection technology with social understanding to analyze, map risks, and determine appropriate interventions. This ensures that every decision is not only data-driven but also socially and culturally relevant within the context of Indonesian society. In daily practice, BSSN consistently refers to and applies core principles of social cybersecurity theory. These principles are not only normative but also serve as concrete guidelines in every stage of operations and decision-making. As explained by the source:

> “When engaging in social cybersecurity efforts, BSSN thoughtfully embraces a range of security principles to safeguard our digital environment. These include raising awareness about operational risks, managing data privacy with care, fostering trust and verifying identities, ensuring secure communication channels, promoting ethical behavior online, and continuously striving to improve through regular evaluation and ongoing development.” (M.A., 21 June 2024).

This affirmation demonstrates that the management of social cybersecurity at BSSN is executed systematically and adaptively rather than sporadically. BSSN consistently guarantees that every digital activity is undertaken with awareness of risks, protection of privacy, trust founded upon verification, secure communication, adherence to digital ethics, and ongoing evaluation to enhance national security standards.

Potential for Approaching a Pseudo-Panopticon

The pseudo-panopticon approach, employing indirect surveillance systems to influence user behavior and prevent the dissemination of disinformation, is acknowledged as an innovative and potentially efficacious cybersecurity strategy. Nonetheless, its implementation presents notable challenges, particularly within the realms of legal, ethical, and individual rights protection. M.A. explicitly articulates this point:

> The pseudo-panopticon approach as a strategy to monitor and control the spread of disinformation offers promising benefits in enhancing security. However, it’s essential to approach this method with careful legal considerations. According to Law Number 17 of 2007 concerning State Intelligence, the government is authorized to carry out intelligence activities,

https://doi.org/10.12928/commicast.v6i2.13673

> including monitoring how information is shared. At the same time, it's essential that any collection and processing of personal data adhere to Law Number 27 of 2022 concerning Personal Data Protection, ensuring that the rights of data subjects are fully respected and protected." (M.A., 21 June 2024).

According to M.A., while this approach can facilitate the management of cybercrimes, adherence to legal provisions introduces the potential risk of infringing upon individual privacy. Therefore, it is imperative to balance security considerations with the protection of personal privacy carefully.

> "While a pseudo-panopticon can be a helpful tool in spotting and managing cybercrimes and the spread of false information, as outlined in the Electronic Information and Transactions Law (ITE), it's important to be mindful of the privacy of individuals. When it comes to copyright, regulated under Law Number 28 of 2014, it's essential to use surveillance technology thoughtfully to respect the rights of creators behind protected works. Overall, although the pseudo-panopticon approach can be quite effective, it's crucial to find the right balance between ensuring security and safeguarding personal privacy, all while sticking to the relevant laws to avoid any misuse or infringement of individual rights." (M.A., 21 June 2024).

Reflections on the statement emphasize that although a pseudo-panopticon can enhance the ability to detect and prevent disinformation, its implementation must be carried out proportionally and within a strict regulatory framework, still limited by principles of personal data protection, individual privacy, and copyright protection. Furthermore, regarding the implementation status, M.A. also affirms that:

"Currently, BSSN has not been given the mandate or authority to implement the pseudo-panopticon. If in the future BSSN is granted the mandate to do so, it will be considered further." (M.A., 21 June 2024).

This statement clarifies that, at present, the pseudo-panopticon remains a concept within BSSN and has not been operationalized, thus requiring legal safeguards and external oversight before the strategy can be widely adopted.

Discussion

Analysis of the social dilemma and BSSN interviews reveals that social media algorithms primarily disseminate misinformation, disinformation, and malinformation. They often display sensational content to boost engagement, fueling polarization, echo chambers, and hindering fact-checking bots (Carley, 2020; Harris, 2019; Wardle & Derakhshan, 2017), automated networks, and advancing AI further complicate the issue by rapidly spreading false content (Hoehe & Thibaut, 2020; Kholili, 2025).

In Indonesia, low levels of digital literacy and deficiencies in algorithmic knowledge contribute to increased public vulnerability to information disorder (Abdillah et al., 2024; Maddock-Ferrie, 2022; Surjatmodjo et al., 2024). BSSN implements initiatives such as systems for monitoring digital media (including social media and news outlets), cross-sector collaboration reporting mechanisms, and digital literacy programs to enhance resilience. These initiatives are based on global guidelines that advocate for multi-sector collaboration and the enhancement of transparency regulations concerning algorithms to combat disinformation (Carley, 2020; Chaudhuri et al., 2025; Roshanaei et al., 2024).

Researchers identify patterns that sustain the dissemination of information disorder driven by social media algorithms. Primarily, there is the pattern where engagement surpasses accuracy, with social media algorithms systematically prioritizing content that elicits emotional responses, sensationalism, or polarization to augment user engagement. This pattern indirectly fosters the proliferation of misinformation and disinformation, as such content is generally more viral and attention-grabbing compared to neutral factual information. This tendency is particularly evident in the widespread dissemination of political hoaxes, SARA issues, and rumors, especially during election periods. Additionally, algorithms reinforce the "filter bubble" or echo chamber phenomenon, whereby users are only exposed to content that corroborates their existing beliefs and preferences. Consequently, polarization and trust in false information intensify, rendering users more susceptible to narrative manipulation. Public opinion becomes susceptible to algorithmic manipulation, thereby complicating efforts to correct misinformation.

Third, field findings and literature studies Akhtar et al., (2023); Alsmadi et al., (2021) substantiate the utilization of bots, automated accounts, and organize networks that extensively generate, amplify, and disseminate information disorder. These activities frequently occur within very short timeframes, surpassing the response capabilities of manual verification systems such as BSSN or government detection mechanisms. Fourth, there exists a knowledge deficiency among the public regarding the functioning of algorithms, with many users unaware of how their digital behaviors (likes, shares, clicks) influence the flow of information they receive. Consequently, individuals tend to trust information that appears frequently on their feeds without verification, thus contributing to the proliferation of information disorder. BSSN addresses this deficiency through digital literacy initiatives and assessments of public understanding. Fifth, emerging threats such as advances in artificial intelligence and content manipulation (e.g., deepfakes, fake news generators) further complicate matters related to information disorder. Sixth, field data and international studies indicate that government regulation remains insufficient, particularly concerning oversight and policies that mandate transparency of algorithms or hold social media platforms accountable in mitigating the spread of misinformation disorder.

Researchers observe that, based on field findings, EMILIE initiated by BSSN can be said to be more than just a slogan; it serves as a strategic framework for strengthening digital literacy that is responsive to the increasingly complex disorder of information. Encouragement and Involvement address Abdillah et al., (2024) research, which states that resilience in digital communities does not solely rely on technology but also public participation and social networks as social filters against disinformation. Measurement and Empowerment align with Carley, (2020) recommendation for continuous monitoring and capacity building so that communities can adapt to the ever-changing social and cyber threats.

Literacy as a pillar of EMILIE is also confirmed by various studies Almatrafi et al., (2024); Kong et al., (2024) indicating that algorithmic literacy and AI understanding are key in efforts to reduce vulnerability to platform algorithm manipulation, as also discussed in the film The Social Dilemma (Iddianto & Azi, 2022). Each stage of EMILIE is designed to be adaptive and needs-based, involving field assessment, the development of specific materials for vulnerable groups, and survey-based evaluation of behavioral changes. This process distinguishes BSSN from traditional one-way digital literacy practices, as it implicitly aligns more closely with the concept of participatory panopticon Zhao Liu, (2025), participatory surveillance akin to China, emphasizing citizen participation as both subjects and objects of oversight.

However, several challenges are faced in implementing EMILIE, such as the volume, speed, and complexity of harmful content. According to Hoehe & Thibaut, (2020) manual detection mechanisms struggle to keep pace with the rapid distribution of misinformation or disinformation driven by algorithms (Harris, 2019; Mulahuwaish et al., 2025). Field data indicates that creating accounts and spreading content can take 10 minutes, but handling the process can take hours.

https://doi.org/10.12928/commicast.v6i2.13673

Additionally, the use of AI for creating deepfakes, automating content, and sophisticated cyberattacks, as identified in the literature Patel et al., (2023), is recognized as a significant challenge by BSSN. Other challenges include resource limitations and platform dependency. Surjatmodjo et al., (2024), highlight that without multi-stakeholder integration, digital literacy and cybersecurity interventions will not be sufficiently effective.

The implementation of EMILIE by BSSN represents a socio-technical system approach Carley, (2020) where technical solutions (automatic detection, monitoring systems) are combined with social interventions (literacy, empowerment, engagement). This aligns with the concept of social cybersecurity, which emphasizes not only protecting digital systems but also behavioral change, cross-sector collaboration, and community empowerment. Theoretically, BSSN's efforts can also be viewed as a form of "soft panopticon," as Foucault described Zhao Liu, (2025), where social surveillance in the digital era shapes citizen behavior without repressive domination. However, privacy concerns, potential misuse, and ethical issues found in studies from China Zhao Liu, (2025) and discussed by Dai et al., (2025), are essential considerations to ensure that the EMILIE model remains transparent and maintains checks and balances.

This research's findings reinforce the literature Carley, (2020); Mulahuwaish et al., (2025); Schmitt & Flechais, (2024), that successful management of the disorder of information requires a multi-layered approach. Strengthening digital literacy based on EMILIE can enhance social resilience against information disorder and algorithm manipulation, but innovations in automatic detection and cross-sector collaboration must support it. Public participation and empowerment also encourage citizens to be active agents in literacy efforts, rather than just targets of intervention. Furthermore, continuous evaluation and adaptation, such as what BSSN practices should be adopted as guiding principles for all national programs, including learning loops from each survey and intervention.

Table 1. Implementation of the EMILIE Framework

Implementations	Findings
Multi-Stakeholder Collaboration	BSSN mobilizes academics, government, industry, and the public (Encouragement and Involvement) for education and reporting on disorder of information.
Evidence-Based and Measurement	Every digital literacy program is evaluated through surveys, KPI monitoring, and statistical impact analysis.
Inclusivity and Local Adaptation	Literacy materials are adapted to the needs assessment of each community, including remote groups; delivery methods are localized.
Empowerment and Knowledge Transfer	Programs not only provide information but also empower the public to become literacy agents within their communities.
Response and Human Resource Challenges	The speed of negative content dissemination outpaces response capacity; limited monitoring personnel; emergence of AI-generated and deepfake content.
Integration of Social and Technical Approaches	Data collection and risk analysis are performed using a combination of technical methods (technology) and social approaches.

The researchers summarized in Table 1, how BSSN's EMILIE Framework promotes cross-sector collaboration, evaluation, adaptation, and community empowerment. Each aspect: Encouragement, Measurement, Involvement, Literacy, and Empowerment, is implemented through digital literacy programs, impact monitoring, and strengthening digital communities. Data shows the framework effectively improves digital literacy, but its long-term success relies on the institution's ability to adapt to AI, content speed, collaboration, and resource challenges, while balancing security, free speech, and privacy.

The panopticon, initially designed by Jeremy Bentham for prisons to allow one guard to see all prisoners without knowing when surveillance occurs, was later used by Michel Foucault as a metaphor for modern disciplinary power rooted in the possibility of being watched, not actual surveillance. Today, it has moved into the digital realm, where algorithms, big data, and AI act as an 'invisible watchtower.' This post-panoptic structure controls, disciplines, and modifies behavior through algorithmic feedback and data-driven normalization, influencing educational and digital governance by encouraging citizens to adjust their actions proactively (Dai et al., 2025).

Studies in China on participatory panopticon Zhao Liu, (2025) show that digital governance lets the state use society as both an object and a surveillance tool. Data technology enhances state control and alters the society-state relationship, increasingly integrating society into digital surveillance. The research warns of power shifts and privacy loss. The quantum panopticon concept Olsson & Ohman, (2025) suggests future risks, with the state potentially storing data long-term to strengthen control and citizens' rights uncertainty. Dai et al., (2025) shows AI surveillance in education normalizes digital monitoring, making individuals more aware of being watched and disciplining themselves. While AI can boost efficiency, it often compromises privacy and independence. Therefore, strict regulation is essential to prevent bias, discrimination, and abuse.

BSSN recognizes the opportunities and challenges of using a pseudo-panopticon as a social cybersecurity strategy. It believes this could enhance disinformation detection by raising public awareness of indirect surveillance. However, BSSN stresses that legal and ethical issues such as human rights, privacy, and data protection must be carefully considered. Currently, BSSN lacks the legal authority to implement this strategy fully. If authorized in the future, transparency, external oversight, and citizen rights must be prioritized to ensure accountable digital surveillance that safeguards civil rights.

Nevertheless, researchers observe that BSSN has implicitly adopted the logic of panopticon surveillance as proposed by Foucault, even without employing specific surveillance devices or programs that might be considered repressive actions. This implementation is evident through strategies such as the EMILIE digital literacy program, which is designed to foster cybersecurity awareness within the community. Through EMILIE, BSSN endeavors to cultivate a digital ecosystem that enhances public awareness of monitoring systems, including regulations and national policies concerning misuse and cybercrime, as well as labels for misinformation or disinformation, warning systems, and content classification. Consequently, this approach encourages self-regulation and the development of behavioral norms within a secure cyberspace. This methodology demonstrates that the efficacy of social cybersecurity strategies need not depend on repression but can be realized through the internalization of norms and collective consciousness, thereby promoting a culture of cybersecurity.

4. Conclusion

The findings of this research hold significant strategic implications for the enhancement of national cyber policies and the advancement of further studies concerning information disturbances. In essence, the novel insights, best practices, and innovations derived from this research can be leveraged by BSSN, relevant agencies, and policymakers to formulate more effective

https://doi.org/10.12928/commicast.v6i2.13673

strategies, regulations, and intervention programs to address the challenges posed by information disturbances in the digital age. The findings also suggest that social media algorithms (Q1), through mechanisms such as recommendations, trending topics, and targeted content, have enabled the rapid and extensive dissemination of malinformation, disinformation, and misinformation in Indonesia. As emphasized in 'The Social Dilemma' and corroborated by existing literature, these algorithms manipulate human psychology, prioritize user engagement over veracity, and create echo chambers that perpetuate harmful content (Akhtar et al., 2023; Harris, 2019). Data collected and documented by BSSN further substantiate that disruptions within this information ecosystem compromise public trust, exacerbate societal polarization, and obstruct initiatives aimed at enhancing digital literacy resilience.

BSSN employs a social cybersecurity strategy founded on EMILIE (Q2), integrating technical detection, rapid response, cross-sector collaboration, adaptive digital literacy, and community empowerment. This framework cultivates collective awareness and digital security norms through education, content classification, and regulatory warnings, rather than through repressive surveillance. Although BSSN has not been delegated a specific mandate to conduct algorithmic oversight via a pseudo-panopticon approach due to legal constraints, researchers highlight that the principles of transparency, external oversight, and the safeguarding of privacy rights and personal data have been implicitly incorporated through the EMILIE framework, which functions as a participatory panoptic governance model, shaping digital citizenship by fostering public awareness and internalizing cybersecurity values rather than through coercion. Currently, EMILIE emphasizes preventive and educational measures, eschewing repressive strategies, and encourages society to regulate digital behavior independently.

Acknowledgment

The author wishes to convey sincere gratitude to all individuals who contributed to the completion of this research. Special thanks are due to the research supervisor for their invaluable guidance and support throughout the process. The author also wishes to acknowledge and thank all those who provided constructive ideas and suggestions that contributed to enriching the study. Deep appreciation is extended to the interviewee from Badan Siber dan Sandi Negara for their assistance and willingness to participate in the research process. Finally, the author extends gratitude to the COMMICAST Journal for facilitating the review, processing, and publication of this work.

References

Abdillah, A., Widianingsih, I., Buchari, R. A., & Nurasa, H. (2024). Big data security &individual (psychological) resilience: A review of social media risks and lessons learned from Indonesia. In *Array*. Elsevier. https://www.sciencedirect.com/science/article/pii/S259000562400002X

Abdo, J. B., Aledhari, M., Qadir, J., Carley, K., & Al-Fuqaha, A. (2025). *A survey of social cybersecurity: Techniques for attack detection, evaluations, challenges, and future prospects*. qspace.qu.edu.qa. https://qspace.qu.edu.qa/handle/10576/65983

Akhtar, M. M., Masood, R., Ikram, M., & Kanhere, S. S. (2023). False Information, Bots and Malicious Campaigns: Demystifying Elements of Social Media Manipulations. *ArXiv eprint arXiv:2308.12497*. https://ui.adsabs.harvard.edu/abs/2023arXiv230812497M/abstract

Almatrafi, O., Johri, A., & Lee, H. (2024). A Systematic Review of AI Literacy Conceptualization. In *Constructs, and*.

Alsmadi, I., Ahmad, K., Nazzal, M., Alam, F., & Al-Fuqaha, A. (2021). Adversarial attacks and defenses for social network text processing applications: Techniques, challenges and future research directions. *ArXiv Preprint ArXiv*. https://arxiv.org/abs/2110.13980

Arroyo, P., Schöttle, A., & Christensen, R. (2021). The ethical and social dilemma of AI uses in the construction industry. In *Proc. 29th Annual Conference of the International Group for Lean Contruction (IGLC)*. academia.edu. https://www.academia.edu/download/97114376/iglc-9ce66acc-87ce-4f66-975b-b8fd667fbd34.pdf

Awadallah, A., Eledlebi, K., & Zemerly, M. J. (2024). Artificial intelligence-based cybersecurity for the metaverse: Research challenges and opportunities. *IEEE Communications Surveys &Tutorials*. https://ieeexplore.ieee.org/abstract/document/10634174/

Carley, K. M. (2020). Social cybersecurity: an emerging science. *Computational and Mathematical Organization Theory*, *26*(4), 365–381. https://doi.org/10.1007/s10588-020-09322-9

Cartwright, E., Chai, Y., & Xue, L. (2025). Leadership in a social dilemma: Does it matter if the leader is pro-social or just says they are pro-social? *Economic Inquiry*. https://doi.org/10.1111/ecin.13256

Chaudhuri, A., Behera, R. K., & Bala, P. K. (2025). Factors impacting cybersecurity transformation: An Industry 5.0 perspective. *Computers &Security*. https://www.sciencedirect.com/science/article/pii/S016740482400573X

Chung, M., & Wihbey, J. (2024). The algorithmic knowledge gap within and between countries: Implications for combatting misinformation. In *Harvard Kennedy School*. misinforeview.hks.harvard.edu. https://misinforeview.hks.harvard.edu/article/the-algorithmic-knowledge-gap-within-and-between-countries-implications-for-combatting-misinformation/

Cloos, J., Greiff, M., & Kempa, K. (2025). The effect of exploiting the public good on climate cooperation: evidence from a collective-risk social dilemma experiment. In *Environment, Development and Sustainability*. Springer. https://doi.org/10.1007/s10668-024-05949-9

Dai, R., Thomas, M. K. E., & Rawolle, S. (2025). Revisiting Foucault's panopticon: how does AI surveillance transform educational norms? *British Journal of Sociology of Education*, *46*(5), 650–668. https://doi.org/10.1080/01425692.2025.2501118

Harris, T. (2019). Our Brains Are No Match for Our Technology. In *International New York Times*. go.gale.com. https://go.gale.com/ps/i.do?id=GALE%7CA608221582&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=22699740&p=AONE&sw=w

Hoehe, M. R., & Thibaut, F. (2020). Going digital: how technology use may influence human brains and behavior. *Dialogues in Clinical Neuroscience*. https://doi.org/10.31887/DCNS.2020.22.2

Husandani, R. A., Utari, P., & Rahmanto, A. N. (2025). Impact of social media disinformation explored in'The Social Dilemma'. *Jurnal ASPIKOM*. http://jurnalaspikom.org/index.php/aspikom/article/view/1534

Iddianto, I., & Azi, R. (2022). Social Effect Of Social Media Revealed In The Social Dilemma Documentary Movie: Post-Truth Perspective. *Seshiski: Southeast Journal of Language and Literary Studies*, *2*(1), 37–50. https://doi.org/10.53922/seshiski.v2i1.3

Kholili, A. (2025). Kultur Digital: Tantangan Dan Peluang Moderasi. In *Kultur Budaya Dan Digital*. repository.iainmadura.ac.id. http://repository.iainmadura.ac.id/1265/2/Layout Kultur Budaya dan Digital.pdf#page=43

Kominfo. (2023). *"Pidato Presiden Jokowi Diduga Menggunakan Bahasa Mandarin."* Kementrian

Komunikasi Dan Informatika. https://www.kominfo.go.id/berita/berita-hoaks/detail/disinformasi-video-pidato-presiden-jokowi-diduga-menggunakan-bahasa-mandarin

Kong, S.-C., Cheung, M.-Y. W., & Tsang, O. (2024). Developing an artificial intelligence literacy framework: Evaluation of a literacy course for senior secondary students using a project-based learning approach. *Computers and Education: Artificial Intelligence*, *6*, 100214. https://doi.org/10.1016/j.caeai.2024.100214

Kurniawan. (2023). *Dokumen Paparan Keamanan Siber Sosial Mewujudkan Media Sosial yang Humanis di Tahun Politik 2024 Guna Menjaga Persatuan Nasional Dalam Rangka Keamanan Nasional*. Kemenkoinfra.Go.Id. https://jdih.kemenkoinfra.go.id/cfind/source/files/perpres/2025/perpres-nomor-12-tahun-2025/lampiran-i-perpres-nomor-12-tahun-2025.pdf

Maddock-Ferrie, B. (2022). A Policy Proposal for Canadian the Government to Counter Disinformation: Countering Disinformation Through Collaboration. *Federalism-E*. https://ojs.library.queensu.ca/index.php/fede/article/view/15368

McDavid, J. (2020). The social dilemma. In *Journal of Religion and Film*. search.proquest.com. https://digitalcommons.unomaha.edu/jrf/vol24/iss1/22

Mujib, M., Fahmi Wardhani, M., & Kurniawan, R. (2023). What Leads To Counterproductive Work Behavior? Predicting The Effect Of Resistance To Change. *Sinergi : Jurnal Ilmiah Ilmu Manajemen*, *13*(2). https://doi.org/10.25139/sng.v13i2.6574

Mulahuwaish, A., Qolomany, B., Gyorick, K., Abdo, J. B., Aledhari, M., Qadir, J., Carley, K., & Al-Fuqaha, A. (2025). A survey of social cybersecurity: Techniques for attack detection, evaluations, challenges, and future prospects. *Computers in Human Behavior Reports*, *18*, 100668. https://doi.org/10.1016/j.chbr.2025.100668

Olsson, E., & Öhman, C. (2025). The Quantum Panopticon: A Theory of Surveillance for the Quantum Era. *Minds and Machines*, *35*(2), 17. https://doi.org/10.1007/s11023-025-09723-2

Patel, Y., Tanwar, S., Gupta, R., Bhattacharya, P., & Davidson, I. (2023). Deepfake generation and detection: Case study and challenges. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10354308/

Roshanaei, M., Khan, M. R., & Sylvester, N. N. (2024). Enhancing cybersecurity through AI and ML: Strategies, challenges, and future directions. In *Journal of Information Security*. scirp.org. https://www.scirp.org/journal/paperinformation?paperid=134347

Schmitt, M., & Flechais, I. (2024). Digital deception: Generative artificial intelligence in social engineering and phishing. In *Artificial Intelligence Review*. Springer. https://doi.org/10.1007/s10462-024-10973-2

Surjatmodjo, D., Unde, A. A., Cangara, H., & Sonni, A. F. (2024). Information Pandemic: A Critical Review of Disinformation Spread on Social Media and Its Implications for State Resilience. *Social Sciences*, *13*(8), 418. https://doi.org/10.3390/socsci13080418

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. tverezo.info. http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf

Zhao Liu, J. (2025). Digital Governance, Dataveillance, and Participatory Panopticon: Public Health Surveillance in China from 2020 to 2022. *Journal of Contemporary China*, 1–17. https://doi.org/10.1080/10670564.2025.2513409