

ARTIFICIAL INTELLIGENCE AND MEDIA INTEGRITY IN DIGITAL MEDIA: A LITERATURE REVIEW ON ALGORITHMIC DISINFORMATION AND COMMUNICATION ETHICS

Riesta Ayu Oktarina^{1,*}, Rahmad Gatut Marwanto², E. Rizky Wulandari³, Edelweis Putri Prima⁴

^{1 2 3 4}Sekolah Tinggi Ilmu Komunikasi Almamater Wartawan Surabaya (Stikosa-AWS)

¹ riestaayu@stikosa-aws.ac.id ; ² rahmadfadhil893@gmail.com ; ³ erizkywulandari@stikosa-aws.ac.id ;

⁴ edelweis.prima@stikosa-aws.ac.id

* corresponding author

ABSTRACT

Article history

Received : November 23, 2025

Revised : December 23, 2025

Accepted : December 27, 2025

Keywords

Artificial Intelligence

Media Integrity

Algorithmic

Disinformation

Communication Ethics

Digital Media

The development of artificial intelligence (AI) has fundamentally transformed the digital media ecosystem, particularly in the processes of information production, distribution, and legitimation. While AI enables efficiency and content personalization, its application also raises serious challenges to media integrity and communication ethics. This article aims to examine the role of artificial intelligence in shaping issues of digital media integrity by employing communication ethics and the political economy of media as the main analytical frameworks. This study adopts a qualitative approach through a literature review of ten international scholarly journals published between 2015 and 2025. The findings indicate that AI operates as a structural actor within the digital media ecosystem by reinforcing the logic of the attention economy, accelerating algorithmic disinformation, and reshaping journalistic ethical practices. At the same time, AI also presents potential as a tool for mitigating disinformation when supported by clear ethical frameworks, algorithmic transparency, and accountable media governance. This article argues that digital media integrity in the age of artificial intelligence constitutes a systemic issue that requires a multidimensional approach integrating communication ethics, technology regulation, and platform accountability.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. Introduction

The development of artificial intelligence (AI) has become one of the most influential factors in the transformation of contemporary communication and digital media. AI no longer functions merely as a technical support tool; rather, it has become structurally integrated into the processes of production, selection, distribution, and legitimation of public information. Through complex algorithmic systems, digital platforms are able to regulate information flows based on user preferences, interaction histories, and audience engagement indicators, thereby

shaping highly personalized media consumption experiences (Floridi, 2020; Zuboff, 2019). This transformation signifies a fundamental shift in communication practices, in which technology plays an active role in determining what is considered relevant, important, and worthy of public consumption.

On the one hand, the implementation of AI in digital media offers efficiency, speed, and levels of information personalization that were previously difficult to achieve. On the other hand, the dominance of algorithmic systems raises serious concerns regarding media integrity and communication ethics. AI-based recommendation systems are generally designed to maximize audience engagement, often prioritizing sensational, emotional, and controversial content over information that is accurate and balanced (Vaccari & Chadwick, 2020). This orientation encourages digital media organizations to adapt their information presentation strategies to algorithmic logic, which may ultimately undermine core journalistic values such as accuracy, verification, and social responsibility.

Issues of media integrity in the digital era have become increasingly complex with the rise of disinformation and information manipulation practices. Unlike the era of conventional media, the dissemination of problematic information today does not rely solely on human decision-making but is also reinforced by technological systems that operate automatically, massively, and continuously. Algorithms and AI contribute to accelerating content circulation, expanding the reach of particular narratives, and creating echo chambers that narrow the diversity of public perspectives (Zuboff, 2019). This condition contributes to declining public trust in the media and weakens the media's function as a social institution that is expected to provide credible and accountable information.

In this context, media integrity can no longer be understood solely as an issue of individual ethics among journalists or editorial teams. Media integrity is also shaped by technological structures that govern the production and distribution of information. AI and platform algorithms function as non-human actors that indirectly shape communication practices, determine information priorities, and influence public perceptions of social reality (Floridi, 2020). Consequently, the crisis of media integrity in the digital era is structural and systemic in nature, involving complex relationships among human actors, technological systems, and the economic interests of digital platforms.

A number of previous studies have addressed issues related to journalistic ethics, information quality, and disinformation in digital media. For example, Qorib et al. (2024) demonstrate that clickbait practices are employed as a strategy to increase audience engagement in online media; however, such practices often result in the neglect of journalistic ethical standards and a decline in information reliability. These findings highlight the tension between the economic interests of digital media and the ethical responsibility of the media to safeguard the quality of public information. Other studies have also emphasized that sensational content tends to gain higher visibility within algorithmic systems compared to content that is informative and in-depth (Vaccari & Chadwick, 2020).

Nevertheless, much of the existing literature still positions technology as an external factor influencing media practices, rather than as a structural actor that actively shapes the communication ecosystem. In contemporary digital media practices, however, algorithms and AI not only facilitate information distribution but also determine the operational logic of media organizations and patterns of audience consumption. Media communication strategies in the digital era are increasingly aligned with platform characteristics and prevailing algorithmic mechanisms. Wulandari (2021) explains that media organizations utilize digital

platforms and social media as communication strategies to build visibility, image, and audience reach. While this strategy is communicatively effective, it also has the potential to intensify message distortion when information framing and distribution are subordinated to algorithmic logic.

Therefore, studies on media integrity in the era of artificial intelligence require a more comprehensive and critical approach. Media integrity must be understood as the outcome of interactions among communication ethics, journalistic practices, and digital technological structures. AI should not be treated as a neutral technology; rather, it should be recognized as a system that embodies specific values, interests, and normative implications within the public sphere (Floridi, 2020). Without an adequate ethical framework, the implementation of AI risks further deepening the crisis of public trust in the media.

Based on the foregoing discussion, this article aims to critically examine the role of artificial intelligence in shaping the crisis of digital media integrity through a literature review approach. This article positions AI as a non-human actor that plays a structural role within the digital communication ecosystem, while also highlighting the ethical and communicative implications of algorithm-based media practices. Accordingly, this study is expected to make a conceptual contribution to the development of communication studies, particularly in advancing the understanding of communication ethics challenges and the sustainability of media integrity in the algorithmic era.

Although this study draws on international scholarly literature, the discussion is particularly relevant to digital media ecosystems in developing countries, where rapid technological adoption is often not accompanied by adequate ethical frameworks and regulatory capacity.

2. Method

This study employs a descriptive qualitative approach using a literature review method to comprehensively examine the influence of artificial intelligence (AI) on digital media integrity. This approach was selected because the research does not aim to test hypotheses or measure variables quantitatively, but rather to analyze patterns, tendencies, and ethical challenges arising from the use of AI within the digital media ecosystem as reflected in existing scholarly literature.

Types and Sources of Data

The data used in this study consist of secondary data obtained from reputable international scholarly journal articles and academic conference proceedings relevant to the research topic. The primary literature analyzed comprises ten international journal articles published between 2015 and 2025. The selection of this time frame was intended to capture recent developments in AI applications in digital media while also tracing the evolution of integrity-related challenges over the past decade.

The decision to focus on ten articles was based on analytical relevance rather than numerical representativeness. During the screening process, articles with overlapping analytical focus or repetitive findings were excluded to avoid redundancy. The final selection of ten articles was considered sufficient to provide conceptual breadth and analytical depth in addressing the research objectives.

The literature sources were retrieved from academic databases including Google Scholar, Cambridge Core, SAGE Journals, SpringerLink, MDPI, IEEE Xplore, and arXiv. All sources included in the analysis are peer-reviewed publications and available in full-text open-access format to ensure transparency and scholarly credibility.

Literature Selection Criteria

The selection of journal articles was guided by the following inclusion criteria:

- (1) the article examines the relationship between artificial intelligence and digital media;
- (2) it addresses issues related to media integrity, disinformation, deepfakes, algorithmic bias, communication ethics, or AI governance;
- (3) it is published in a reputable scholarly journal and has undergone a peer-review process; and
- (4) it demonstrates conceptual relevance to communication and media studies.

Publications categorized as non-academic opinions, popular reports, or studies lacking methodological clarity were excluded from the analysis.

Data Collection Techniques

Data collection was conducted through a structured literature search using predefined keywords, including: *“Artificial Intelligence and Media Integrity,” “AI and Disinformation,” “Deepfake in Journalism,” “Algorithmic Bias in News,”* and *“Ethics of AI in Digital Media.”* Each article meeting the inclusion criteria was documented based on its metadata, including title, author(s), year of publication, journal name, and Digital Object Identifier (DOI), prior to further analysis.

Data Analysis Techniques

Data analysis was carried out using qualitative thematic analysis involving three main stages. First, thematic identification was conducted by categorizing the selected articles into key analytical themes such as AI-driven disinformation, deepfake technologies, algorithmic bias, and ethical governance of digital media. Second, manual coding was applied to record the research objectives, methodological approaches, principal findings, and ethical implications of each article. Third, synthesis of findings was undertaken by comparing similarities, differences, and scholarly contributions across the reviewed studies in explaining the relationship between artificial intelligence and digital media integrity.

Thematic saturation was considered achieved when subsequent articles no longer generated new analytical categories or conceptual insights relevant to the research questions. The recurring emergence of core themes across the reviewed literature indicated that additional sources would likely reinforce existing patterns rather than produce substantively new findings.

Methodological Framework Alignment

Although this study does not adopt a fully systematic review protocol, the methodological procedure aligns with a semi-systematic literature review framework. The review applied transparent inclusion criteria, a defined temporal scope, systematic keyword searches across academic databases, and thematic synthesis of findings. This approach allows

analytical flexibility while maintaining methodological rigor appropriate for exploratory and conceptual research in communication studies.

Validity and Trustworthiness of Data

To ensure validity and trustworthiness, this study employed source triangulation by comparing findings from multiple journals addressing similar themes across different contexts and perspectives. Only peer-reviewed articles were included to minimize bias and enhance credibility. All references were managed using citation management software to ensure consistency and accuracy in citation practices.

3. Results and Discussion

Artificial Intelligence as a Structural Actor in Digital Media

The review of ten international journal articles published between 2015 and 2025 indicates that artificial intelligence has undergone a significant shift in its role within the digital media ecosystem. AI no longer functions merely as a technical support tool in the news production process; instead, it has emerged as a structural actor that influences editorial decision-making, content curation, and the distribution of information to the public. Through automated journalism and algorithm-based recommendation systems, AI has assumed functions that were previously under the control of human editorial teams (Braun & Eklund, 2023).

This structural perspective differs from techno-deterministic views that frame artificial intelligence merely as an autonomous force shaping media outcomes, as well as from instrumental approaches that treat AI as a neutral tool fully controlled by human actors. Rather than positioning AI as either an independent determinant or a passive instrument, this study aligns with a socio-technical understanding in which artificial intelligence operates within institutional, economic, and normative structures. From this viewpoint, AI shapes media practices while simultaneously reflecting organizational goals, platform logics, and market-driven imperatives. Such a position allows for a more nuanced analysis of media integrity by acknowledging both technological agency and human responsibility within algorithmic systems.

Algorithmic Disinformation and the Logic of the Attention Economy

Key findings from the reviewed literature indicate that AI intensifies the dissemination of disinformation through algorithmic mechanisms oriented toward audience engagement. Social media algorithms and digital news platforms tend to prioritize content with high click-through and interaction potential, without adequately considering the quality and accuracy of information (Floridi & Cowls, 2021). As a result, disinformation not only spreads more rapidly but also gains a form of symbolic legitimacy, as it is continuously amplified and recommended by algorithmic systems.

Empirical studies across different platforms illustrate how attention-based algorithms intensify the circulation of disinformation. Comparative findings show that sensational or emotionally charged content is more likely to be amplified by recommendation systems on social media platforms than fact-based reporting, regardless of national context. For example, studies conducted in both Western and Global South media environments indicate similar patterns in which algorithmic visibility favors engagement-driven narratives over informational quality. These cross-contextual patterns suggest that algorithmic

disinformation is not confined to specific regions, but represents a systemic outcome of attention-oriented media architectures.

Clickbait practices represent a concrete manifestation of the degradation of media integrity within the logic of the attention economy. Qorib et al. (2024) demonstrate that sensational headlines are strategically employed to increase audience engagement, yet such practices directly contribute to the erosion of journalistic ethical standards and the decline of information reliability. When combined with AI-driven algorithms, disinformation ceases to be incidental and instead becomes embedded within the systemic processes of digital media production and distribution.

Beyond text-based and narrative disinformation, advances in artificial intelligence have also introduced new challenges in the form of visual and audiovisual manipulation. At this point, issues of media integrity extend beyond the veracity of information to encompass the reliability of visual evidence, which has long served as a foundational source of public trust in the media.

Deepfakes and the Crisis of Visual Authenticity

The findings of this review further affirm that deepfake technology poses a serious threat to the integrity of digital media. Deepfakes enable highly realistic visual and audiovisual manipulation that is increasingly difficult to distinguish from authentic content (Verdoliva, 2020). Research by Vaccari and Chadwick (2025) demonstrates that AI-generated visual content has a significant impact on public perception, particularly in political contexts and strategic public issues.

This crisis of visual authenticity has shifted the paradigm of trust in visual evidence within journalism. Whereas visual content was previously positioned as a guarantor of truth, it has now become a source of epistemological uncertainty. Cases of deepfake usage in electoral contexts in India and the United States illustrate how visual manipulation can shape public perceptions before verification or clarification processes can take place (Zhou et al., 2020). Such conditions undermine the media's role as a provider of verifiable and reliable information.

It is important to distinguish between the technical challenges posed by deepfake technologies and their broader ethical consequences. Technically, deepfakes challenge existing verification mechanisms by producing highly realistic synthetic visual content that is difficult to detect using conventional forensic tools. Ethically, however, the implications extend beyond detection difficulties to issues of trust, accountability, and public deception. While technical solutions may reduce the circulation of manipulated content, they do not fully address the normative question of responsibility when visual evidence itself becomes epistemologically unstable within digital media ecosystems.

The crisis of visual authenticity triggered by deepfake technology indicates that issues of media integrity do not operate in isolation but are closely intertwined with the ways algorithms filter and recommend information. In this context, algorithmic bias emerges as a crucial factor that intensifies informational polarization and constrains the space for public dialogue.

Algorithmic Bias and Information Polarization

The reviewed literature also reveals that AI systems operate on historically situated and non-neutral data, which makes them prone to reproducing and reinforcing social,

political, and ideological biases (Shorey, 2023b). AI-based recommendation systems tend to amplify user preferences and generate echo chambers, in which audiences are predominantly exposed to information that aligns with their existing viewpoints (Braun & Eklund, 2023).

The impact of algorithmic bias is reflected in the intensification of information polarization and the weakening of the media's function as a space for public dialogue. AI does not merely filter information; it actively participates in shaping the social reality consumed by the public. From the perspective of media integrity, this condition poses a serious challenge to the principles of balance and diversity of perspectives.

As algorithmic bias and information polarization continue to intensify, fundamental questions arise regarding the extent to which oversight and regulatory mechanisms are capable of keeping pace with technological developments. This situation highlights a growing gap between the rapid pace of AI innovation and the readiness of ethical frameworks and digital media governance structures.

From an audience-centered perspective, algorithmic bias not only structures information exposure but also shapes patterns of media consumption and interpretation. Research on audience behavior suggests that users often remain unaware of algorithmic filtering processes, perceiving personalized content streams as neutral representations of reality. This limited algorithmic awareness reinforces selective exposure and confirmation bias, thereby intensifying information polarization. Consequently, media integrity challenges emerge not only at the level of content production but also through audience interaction with algorithmically curated information environments.

Regulatory Gaps and Ethical Challenges in Digital Media

The findings of this review indicate a significant lag between the rapid development of AI technologies and the existing regulatory and ethical frameworks governing digital media. Most of the reviewed studies highlight the weakness of AI governance in the media sector, particularly with regard to algorithmic transparency and technological accountability (Giansanti, 2022b). Digital platforms tend to obscure their content recommendation mechanisms under the justification of technological protection, thereby limiting public and regulatory access to the evaluation of potential bias and information manipulation.

In developing countries, including Indonesia, these challenges are further exacerbated by low levels of public digital literacy and the limited capacity of media oversight institutions. As a result, the public remains in a vulnerable position, exposed to information generated or filtered by AI systems without adequate mechanisms of control and accountability.

Amid regulatory limitations and ongoing ethical challenges, some strands of the literature view artificial intelligence as part of the solution to media integrity problems. However, the use of AI as a tool for mitigating disinformation introduces a new paradox that warrants critical examination.

Regulatory responses to artificial intelligence in digital media vary across regions. In the European context, emerging regulatory frameworks emphasize transparency, accountability, and risk-based governance of AI systems, particularly in relation to public communication and information integrity. In contrast, many developing countries face regulatory gaps due to rapid technological adoption that outpaces institutional capacity and policy development. This comparison highlights that challenges to media integrity are not solely technological, but are also shaped by regional differences in regulatory maturity and governance models.

AI as a Solution and the Ethical Paradox

Interestingly, the literature review also indicates that AI holds significant potential as a tool for mitigating disinformation. Studies by Ahmed (2025) and Wang and Kagal (2024) demonstrate that natural language processing–based models are capable of detecting hoaxes rapidly and efficiently, particularly in political and public health contexts. AI also has the potential to assist journalists in real-time information verification processes.

However, this potential is inherently paradoxical. The AI systems used to detect disinformation often operate within the same technological infrastructures as those that accelerate the dissemination of problematic content. Without algorithmic transparency and robust ethical oversight, the use of AI as a solution risks reinforcing informational power asymmetries between digital platforms and the public.

To understand this paradox more deeply, it is essential to examine how artificial intelligence functions within data-driven decision-making processes. This perspective helps explain why AI can simultaneously serve as a mechanism of control and a source of risk within the digital media ecosystem.

Synthesis of Findings: Media Integrity as a Systemic Issue

Based on the overall results and discussion, it can be concluded that the crisis of digital media integrity in the era of artificial intelligence is systemic in nature. Media integrity is no longer determined solely by individual journalistic ethics, but is also shaped by technological design, algorithmic logics, and the economic interests of digital platforms. AI functions as an accelerator that can amplify both ethical and unethical communication practices, depending on the governance frameworks within which it operates.

These findings underscore that efforts to safeguard digital media integrity require a multidimensional approach that integrates communication ethics, media literacy, technological regulation, and platform accountability. Without such an approach, artificial intelligence is likely to continue exacerbating the crisis of public trust in digital media.

AI-Based Decision-Making Logic and Its Implications for Digital Media

Additional literature indicates that artificial intelligence operates through large-scale data-driven decision-making logic. Nugroho et al. (2025) explain that AI is capable of identifying hidden patterns, predicting trends, and generating rapid and accurate decision recommendations through big data processing. In the context of digital media, these capabilities are translated into content recommendation systems, news distribution optimization, and the prioritization of issues based on audience engagement metrics.

However, when this decision-making logic is applied within the media industry, editorial decisions are no longer fully grounded in journalistic ethical considerations, but increasingly shaped by algorithmic efficiency and profit potential. This condition reinforces the finding that AI functions as a non-human decision-maker influencing the direction of public discourse. Consequently, media integrity undergoes a shift from normative values toward an instrumental, data-driven logic.

The dominance of AI in decision-making processes underscores the ambivalent nature of this technology. On the one hand, AI offers efficiency and accuracy; on the other hand, it risks displacing the human dimension within communication and journalistic practices.

AI as a Double-Edged Sword: Technological Efficiency and the Risk of Dehumanization

A number of studies emphasize that the implementation of AI yields positive impacts in the form of increased efficiency, productivity, and speed in decision-making processes. Across various sectors, including business, education, and public services, AI is capable of replacing repetitive tasks and accelerating analytical procedures. However, the same body of literature also cautions that the growing dominance of AI may lead to dehumanization, defined as the diminishing role of human values in social processes.

In the context of digital media, the risk of dehumanization becomes evident when journalism is reduced to a mere data commodity. Algorithms lack moral sensitivity, empathy, and contextual awareness—qualities that are inherent to human journalists. As a result, decisions generated by AI, while efficient, may overlook the social, psychological, and ethical consequences of information dissemination. These findings reinforce the argument that technological efficiency does not necessarily align with the quality of public communication.

Beyond the risk of dehumanization, the literature also indicates that the widespread adoption of AI contributes to the emergence of new inequalities in information distribution. Such inequalities reflect increasingly asymmetrical power relations within the digital media ecosystem.

AI, Productivity, and Information Inequality in the Media Ecosystem

Duila's (2023) study demonstrates that the widespread adoption of AI indeed enhances productivity, while simultaneously generating new forms of inequality in both labor structures and information distribution. In digital media, this inequality is evident in the dominance of large platforms that possess access to advanced AI technologies, whereas smaller and independent media organizations lag behind in terms of technological capacity and visibility.

Such inequalities have direct implications for media integrity, as the production and distribution of information become increasingly concentrated among a limited number of actors. In this context, AI not only improves productivity but also reinforces power relations within the digital public sphere. Media outlets that are unable to adapt to algorithmic logics risk losing audience attention, even when the content they produce adheres to ethical standards and journalistic quality.

Ethical, Humanistic, and Social Responsibility Dimensions of AI

The literature on technology ethics emphasizes that the development and deployment of AI must take into account humanistic dimensions and social responsibility. AI cannot be treated as a neutral system, as it reflects the values, interests, and biases of those who design and operate it. In the media context, this means that algorithms actively shape the social reality consumed by the public.

A humanistic approach requires that AI in media should not fully replace human judgment, but rather function as an assistive tool operating under the principles of communication ethics. In the absence of a clear ethical framework, AI risks deepening public alienation, reducing the quality of democratic deliberation, and weakening the media's role as a space for social dialogue.

To formulate a more applicable ethical approach, lessons from other sectors become highly relevant. Studies on the use of AI in education and social domains indicate that technology is most effective when governed by human values and responsibility—an essential principle that should also be applied in digital media practices.

Lessons from the Education and Social Sectors: Relevance for Digital Media

Studies on the application of AI in the education sector indicate that AI can enhance learning effectiveness through content personalization and adaptive systems, while still requiring the role of educators as guides of values and contextual understanding. These findings are highly relevant to digital media, where AI may be utilized to personalize information delivery but should not replace ethical curation and the social responsibility of media institutions.

In other words, lessons from the education and social sectors emphasize that AI is most effective when positioned as a co-actor rather than as a sole decision-maker. This perspective reinforces the argument that digital media integrity can only be maintained when the use of AI is accompanied by human oversight, ethical regulation, and a strong orientation toward the public interest.

4. Conclusion

This study concludes that artificial intelligence has become a structural element that significantly shapes the integrity of digital media. AI not only influences the efficiency of information production and distribution, but also determines the visibility, legitimacy, and meaning of public information through algorithmic mechanisms. The consequences of this condition include the emergence of systemic ethical issues, such as algorithmic disinformation, visual manipulation, biased information distribution, and the erosion of journalistic principles oriented toward the public interest. In this context, media integrity can no longer be understood solely as the responsibility of individual media actors, but rather as the outcome of interactions among technology, journalistic practices, and the economic interests of digital platforms.

Based on these findings, future research should be directed toward empirical studies that examine more deeply how artificial intelligence is operationalized within institutional media practices, particularly in relation to editorial decision-making and algorithmic governance. Further studies are also needed to explore audience perceptions and responses to AI-curated content, as well as the effectiveness of ethical frameworks and regulatory mechanisms across diverse social and cultural contexts. Interdisciplinary approaches that integrate communication studies, technology ethics, and media policy are expected to enrich the understanding of strategies for safeguarding digital media integrity in the era of artificial intelligence.

Acknowledgment

The authors would like to express their gratitude to all parties who contributed to the completion of this article. Special appreciation is extended to colleagues and academic peers who provided valuable insights, feedback, and discussions that enriched the development of this study. The authors also acknowledge the support of the academic environment that facilitated access to relevant literature and scholarly resources, enabling this research to be conducted effectively.

References

- Braun, J. A., & Eklund, J. L. (2023). Fake news, real money: Ad tech platforms, profit-driven hoaxes, and the business of journalism. *Digital Journalism*, 11(2), 176–195.
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- Duila, M. (2023). Dampak artificial intelligence terhadap produktivitas dan ketimpangan kerja digital. *Jurnal Ilmu Sosial dan Humaniora*, 12(2), 251–260.
- Fahrudin, A., Prasetyo, A., & Lestari, N. (2022). Etika dan tantangan kecerdasan buatan dalam masyarakat digital. *Jurnal Pengembangan Ilmu Manajemen*, 10(1), 45–57.
- Floridi, L. (2020). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- Floridi, L., & Cowls, J. (2021). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 3(1).
- Giansanti, D. (2022). The regulation of artificial intelligence: Ethical and governance issues. *Healthcare*, 10(6), 1–15.
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- McQuail, D. (2010). *McQuail's mass communication theory* (6th ed.). SAGE Publications.
- Qorib, F., Oktarina, R. A., Mardiana, S., & Wulandari, E. R. (2024). Clickbait in the era of information technology: Evaluating engagement strategies and ethical journalism in online media. In *Proceedings of the Seventh International Conference on Vocational Education and Electrical Engineering (ICVEE)* (pp. 204–209). IEEE.
- Shorey, S. (2023). Algorithmic bias and the limits of automation. *AI & Society*, 38(2), 589–601.
- Sutojo, T., Mulyanto, E., & Suhartono. (2011). *Kecerdasan buatan*. Andi Offset.
- 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.
- Verdoliva, L. (2020). Media forensics and deepfakes: An overview. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 910–932.
- Wang, Y., & Kagal, L. (2024). Artificial intelligence for misinformation detection: Challenges and opportunities. *Journal of Artificial Intelligence Research*, 78, 1–26.
- Wulandari, E. R. (2021). Digital strategy of political communication media on social media. *Wasilatuna: Jurnal Komunikasi dan Penyiaran Islam*, 4(1), 28–42.
- Zhou, Q., Yang, X., Lim, H., & Li, J. (2020). Deepfake detection: A survey. *IEEE Access*, 8, 128266–128282.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.