https://research.adra.ac.id /index.php/jssut/
P - ISSN: 3026-5959
E - ISSN: 3026-605X

Beyond Fake News: The Rise of Deepfakes and the Crisis of Epistemic Trust in Digital Journalism

Prasojo[1], Ishara Ratnayake[2], Kumudu Wijerathna[3]

[1] Universitas Bhayangkara Jakarta Raya, Indonesia
[2] Sri Jayewardenepura University, Sri Lanka
[3] Sri Jayewardenepura University, Sri Lanka

ABSTRACT

Background. The rapid advancement of artificial intelligence has transformed the digital information landscape, enabling the creation of highly realistic synthetic media known as deepfakes. Unlike conventional fake news, deepfakes possess the capacity to manipulate audiovisual evidence in ways that challenge public perceptions of authenticity and truth. The growing proliferation of deepfake content has raised significant concerns regarding the credibility of digital journalism, the reliability of information sources, and the erosion of epistemic trust within contemporary communication environments.
Purpose. This study aims to examine how the rise of deepfakes contributes to the crisis of epistemic trust in digital journalism and to explore the broader implications of synthetic media for journalistic credibility and public knowledge.
Method. A qualitative research design based on a critical literature review and thematic content analysis was employed. Data were collected from peer-reviewed journal articles, institutional reports, policy documents, and documented case studies published between 2018 and 2025.
Results. Findings reveal that deepfakes undermine trust not only by disseminating deceptive content but also by fostering uncertainty regarding the authenticity of legitimate information. Authenticity ambiguity, verification challenges, audience skepticism, and institutional trust erosion emerged as the dominant themes across the analyzed literature.
Conclusion. The study concludes that deepfakes represent a fundamental epistemic challenge that extends beyond misinformation. Preserving trust in digital information ecosystems requires integrated responses involving technological innovation, journalistic verification, media literacy, and governance frameworks.

KEYWORDS

Deepfakes, Digital Journalism, Epistemic Trust, Media Credibility, Synthetic Media.

Citation: Prasojo, Prasojo., Ratnayake, I., & Wijerathna, K. (2026). Beyond Fake News: The Rise of Deepfakes and the Crisis of Epistemic Trust in Digital Journalism. *Journal of Social Science Utilizing Technology*, *4*(2), 173–188. https://doi.org/10.70177/jssut.v4i2.3933

Correspondence:

Prasojo,
prasojo@dsn.ubharajaya.ac.id

Received: October 9, 2025

Accepted: March 14, 2026

Published: April 30, 2026

INTRODUCTION

Digital transformation has fundamentally altered the production, distribution, and consumption of news across contemporary societies. Expansion of social media platforms, algorithm-driven information ecosystems, and artificial intelligence technologies has accelerated the speed at which information circulates globally. Journalism, once positioned as a primary gatekeeper of public information, now operates within a highly decentralized communication environment where content can be created, modified, and disseminated by a wide range of actors Such

developments have generated unprecedented opportunities for information accessibility while simultaneously introducing new vulnerabilities to information integrity and public trust.

Artificial intelligence has recently enabled the emergence of deepfakes, a form of synthetic media capable of generating highly realistic audio, video, and visual content that closely mimics real individuals, events, and situations (Allen dkk., 2025). Advances in generative adversarial networks and other machine learning techniques have significantly improved the quality and accessibility of deepfake production (Fallis, 2025). Distinctions between authentic and manipulated content have become increasingly difficult for ordinary audiences to recognize, creating serious concerns regarding misinformation, disinformation, media manipulation, and digital deception.

Epistemic trust, defined as the confidence individuals place in information sources and their ability to convey truthful representations of reality, has become a central concern within digital journalism (Wirtz dkk., 2026). Public trust in news institutions already faces significant challenges due to political polarization, information overload, and the proliferation of fake news (Vineela dkk., 2025). Deepfake technologies introduce a more complex threat because they undermine not only the credibility of specific information but also the broader capacity of audiences to distinguish truth from fabrication (Muralidharan, 2025). Growing uncertainty regarding the authenticity of digital content has elevated deepfakes from a technological issue to a critical societal and journalistic challenge.

Deepfakes present a significant challenge to the foundational principles of journalism, which depend upon accuracy, verification, transparency, and accountability (Díaz dkk., 2026). Journalistic institutions increasingly encounter situations in which manipulated audiovisual content can rapidly circulate before verification processes are completed (Yazdi dkk., 2026). False narratives supported by convincing synthetic media possess the potential to influence public opinion, distort political discourse, damage reputations, and shape collective perceptions of reality (Cazzamatta, 2025). Existing journalistic practices are therefore being tested by technologies capable of producing sophisticated forms of deception.

Crisis of epistemic trust emerges when audiences become uncertain about the authenticity of information regardless of whether specific content is true or false (Bahra & Fennan, 2025). Deepfakes contribute to this phenomenon by creating an environment in which visual and audio evidence can no longer be automatically regarded as reliable indicators of reality (Ramos-Zaga, 2025a). Individuals exposed to repeated instances of manipulated media may develop skepticism toward legitimate journalistic reporting, leading to broader erosion of trust in news organizations and public communication systems (Fuchs, 2025). Such conditions create challenges that extend beyond misinformation and enter the realm of epistemological uncertainty.

Regulatory frameworks, technological detection tools, and journalistic verification mechanisms have struggled to keep pace with the rapid evolution of synthetic media technologies (Ramos-Zaga, 2025b). Existing approaches often focus on identifying fake content after dissemination rather than addressing the broader implications for public trust and democratic communication (Karanja & Nzau, 2026). Insufficient understanding of how deepfakes influence epistemic trust within digital journalism limits the ability of scholars, journalists, and policymakers to develop effective responses (Imran dkk., 2025). This situation highlights the necessity of examining deepfakes as both a technological and socio-epistemic problem.

This study aims to investigate the relationship between the rise of deepfake technologies and the emerging crisis of epistemic trust within digital journalism (Teixeira, 2025). Particular attention is directed toward understanding how synthetic media challenges traditional assumptions regarding authenticity, credibility, and truthfulness in journalistic communication (Wanjari dkk., 2026).

Examination of these issues contributes to a deeper understanding of the evolving dynamics between technology and information integrity in contemporary media environments.

Analysis of the mechanisms through which deepfakes influence audience perceptions constitutes another important objective of the study (Anani-Bossman, 2026). Exploration of how individuals evaluate, interpret, and respond to potentially manipulated content provides insights into changing patterns of trust formation in digital information ecosystems (Hynek dkk., 2025). Understanding these processes is essential for assessing the broader implications of synthetic media on public knowledge and democratic participation.

Development of a conceptual framework for understanding epistemic trust in the age of artificial intelligence represents a further objective of this research (Zhang dkk., 2025). Findings are expected to contribute to scholarly discussions concerning digital journalism, media credibility, information disorder, and technological governance (McBrayer, 2026). Recommendations derived from the study may assist journalists, media organizations, educators, and policymakers in addressing the challenges posed by increasingly sophisticated forms of synthetic media.

Existing scholarship has extensively examined fake news, misinformation, and disinformation within digital media environments (Abraham dkk., 2025). Numerous studies have explored the causes, dissemination patterns, and societal consequences of false information circulating through online platforms (Doan, 2026). Research has also investigated the role of social media algorithms, political polarization, and cognitive biases in shaping public vulnerability to misinformation (Dwyer, 2024). Despite these contributions, many studies continue to focus primarily on textual misinformation rather than synthetic audiovisual manipulation.

Emerging literature on deepfakes has largely concentrated on technical dimensions such as detection algorithms, machine learning architectures, cybersecurity concerns, and content authentication mechanisms (Saranya dkk., 2026). Although these studies provide valuable insights into technological solutions, relatively limited attention has been devoted to the epistemological consequences of deepfakes for journalism and public trust (Yang, 2025). Questions concerning how synthetic media alters audience perceptions of credibility, evidence, and journalistic authority remain insufficiently explored within current academic discourse.

Research examining trust in journalism often treats misinformation, media bias, political polarization, and institutional performance as primary explanatory factors (B dkk., 2025). Deepfakes introduce a distinct challenge because they undermine confidence in the evidentiary foundations upon which journalism traditionally relies (Murugalakshmi dkk., 2025). Limited integration between deepfake studies and theories of epistemic trust creates an important gap in the literature (Dinana dkk., 2025). Absence of comprehensive analyses connecting synthetic media technologies with broader crises of knowledge verification highlights the need for further investigation.

Novelty of this study lies in its conceptual repositioning of deepfakes as a catalyst for an epistemic crisis rather than merely a new form of misinformation (Kim, 2026). Existing research frequently frames deepfakes as technological threats requiring technical detection solutions (Whitelaw, 2026). This study advances a broader perspective by examining how synthetic media challenges fundamental assumptions about truth, evidence, credibility, and knowledge production within digital journalism (Hoffmann dkk., 2025). Such an approach extends beyond content authenticity and addresses deeper questions concerning public trust and information legitimacy.

The study introduces an interdisciplinary analytical framework that integrates perspectives from journalism studies, communication theory, media ethics, epistemology, and artificial intelligence governance (Ismail dkk., 2025). Combining these perspectives enables a more

comprehensive understanding of the societal implications of deepfake technologies (Kassymbekova dkk., 2025). Interdisciplinary examination facilitates exploration of how technological developments intersect with cognitive, institutional, and democratic dimensions of information trust.

Significance of this research is grounded in the increasing prevalence of generative artificial intelligence and its potential impact on democratic societies. Journalistic institutions play a critical role in facilitating informed public discourse and supporting democratic decision-making (Hameed & Khudhair, 2026). Erosion of epistemic trust threatens the effectiveness of these functions by weakening confidence in reliable sources of information (Valenzuela dkk., 2026). Findings from this study are expected to provide theoretical contributions to scholarship on digital journalism while offering practical insights for media organizations, policymakers, educators, and technology developers seeking to strengthen trust in contemporary information ecosystems.

RESEARCH METHODOLOGY

This study employs a qualitative research design based on a critical literature review and thematic content analysis to investigate the relationship between deepfake technologies and the crisis of epistemic trust in digital journalism (Avides Moreira dkk., 2026). A qualitative approach is particularly suitable because the study seeks to explore complex social, communicative, and epistemological phenomena that cannot be adequately understood through numerical measurement alone (Cheng & Hsu, 2025). Analytical emphasis is placed on understanding how deepfakes challenge traditional assumptions regarding authenticity, credibility, and trustworthiness in contemporary news environments.

The research is grounded in an interpretive paradigm that views trust as a socially constructed and context-dependent phenomenon. Deepfakes are examined not merely as technological artifacts but as communicative interventions capable of reshaping public perceptions of truth and reality. The study integrates perspectives from journalism studies, media ethics, communication theory, artificial intelligence governance, and epistemology to develop a comprehensive understanding of the challenges posed by synthetic media in digital information ecosystems.

Comparative thematic analysis is employed to identify recurring patterns, conceptual relationships, and emerging concerns within the scholarly and professional discourse surrounding deepfakes and journalism. Examination of multiple sources enables the study to capture diverse viewpoints regarding the implications of synthetic media for journalistic credibility, public trust, democratic communication, and information verification processes. Such an approach facilitates a nuanced interpretation of the evolving relationship between technological innovation and epistemic stability.

The population of this study consists of academic publications, policy reports, journalistic investigations, institutional documents, and professional guidelines addressing deepfakes, digital journalism, misinformation, media credibility, and epistemic trust. These materials represent the primary sources through which contemporary discussions concerning synthetic media and information integrity are articulated and debated. Inclusion of diverse source categories ensures a comprehensive understanding of both theoretical and practical dimensions of the phenomenon under investigation.

Purposive sampling is utilized to select sources that are directly relevant to the objectives of the study. Selection criteria include scholarly significance, thematic relevance, publication credibility, and contribution to discussions concerning deepfakes and trust in digital journalism. Academic articles published in peer-reviewed journals, reports issued by international

organizations, media watchdog institutions, technology research centers, and recognized journalism associations are prioritized during the sampling process.

The final sample consists of scholarly literature and institutional publications produced between 2018 and 2025, a period characterized by significant advancements in generative artificial intelligence and increasing public concern regarding synthetic media. Sources addressing deepfake creation, detection technologies, media manipulation, audience trust, journalistic verification, and democratic communication are included in the analysis. This sampling strategy ensures that the study reflects recent developments and contemporary debates surrounding deepfakes and digital journalism.

The primary instrument employed in this study is a thematic analysis framework developed to systematically examine the content of selected documents. The framework includes analytical categories related to deepfake technology, misinformation, epistemic trust, journalistic credibility, audience perception, verification practices, ethical concerns, and regulatory responses. These categories guide the identification and organization of relevant themes across the collected sources.

A document analysis matrix is utilized to facilitate systematic comparison and interpretation of the selected materials. The matrix records bibliographic information, research focus, conceptual arguments, methodological approaches, key findings, and implications related to deepfakes and trust in journalism. Structured organization of information enhances analytical consistency and supports the identification of recurring patterns and conceptual relationships within the literature.

Coding protocols are also employed to strengthen analytical rigor throughout the research process. Open coding is initially used to identify significant concepts emerging from the data, followed by axial coding to establish connections among related themes. Application of these procedures enables the development of higher-order categories that explain how deepfakes contribute to the erosion of epistemic trust and influence journalistic practices within contemporary digital environments.

Data collection begins with a systematic search of academic databases, institutional repositories, professional journalism platforms, and policy archives. Keywords such as “deepfakes,” “synthetic media,” “digital journalism,” “epistemic trust,” “media credibility,” “misinformation,” and “artificial intelligence” are employed to identify relevant sources. Retrieved materials are subsequently screened according to predetermined inclusion and exclusion criteria to ensure relevance and quality.

Analytical procedures involve several interconnected stages. The first stage focuses on document familiarization through repeated reading and preliminary note-taking. The second stage consists of coding and categorizing relevant information according to the analytical framework. Themes associated with trust formation, media authenticity, technological deception, verification practices, audience responses, and journalistic adaptation are systematically identified and documented. Continuous comparison among sources is conducted to strengthen analytical consistency and thematic coherence.

Interpretation constitutes the final stage of the research process. Identified themes are synthesized and analyzed in relation to broader theoretical discussions concerning epistemic trust, journalism, and digital communication. Findings are evaluated to determine how deepfake technologies influence public confidence in news institutions and information sources. Conclusions are subsequently formulated to provide conceptual insights and practical recommendations regarding the preservation of trust and credibility in an increasingly synthetic media environment shaped by advances in artificial intelligence.

RESULT AND DISCUSSION

Analysis of the selected literature, institutional reports, and journalism studies reveals a substantial increase in scholarly attention toward deepfakes and epistemic trust between 2018 and 2025. A total of 85 documents meeting the inclusion criteria were analyzed, comprising peer-reviewed journal articles, policy reports, media industry publications, and technology governance studies. Content classification identified four dominant thematic categories: deepfake detection and technology (31%), misinformation and disinformation (27%), journalism and media credibility (24%), and epistemic trust and audience perception (18%). Distribution patterns suggest that research initially focused on technical aspects of deepfake production and detection before gradually expanding toward broader societal and journalistic implications.

Comparative examination of the literature demonstrates a growing recognition that deepfakes represent a challenge extending beyond conventional misinformation. Findings indicate that concerns regarding trust erosion appeared in approximately 79% of the analyzed documents, while discussions related to journalism credibility emerged in 72% of the sources. References to democratic implications were identified in 64% of the documents, whereas regulatory concerns appeared in 58%. These trends suggest that scholarly attention has increasingly shifted from technological capabilities toward the societal consequences of synthetic media proliferation.

Table 1. Distribution of Major Themes in Deepfake and Digital Journalism Research (2018–2025)

Research Theme	Frequency (n)	Percentage (%)
Deepfake Technology and Detection	26	31
Misinformation and Disinformation	23	27
Journalism and Media Credibility	20	24
Epistemic Trust and Audience Perception	16	18
Total	85	100

The data indicate that the evolution of deepfake research has followed a trajectory from technological investigation toward broader societal analysis. Early studies concentrated on algorithmic mechanisms, detection systems, and cybersecurity implications. Recent publications increasingly address concerns related to public trust, media authenticity, and democratic communication. This shift reflects growing awareness that deepfakes represent not merely a technical challenge but also an epistemological and institutional problem affecting information ecosystems.

Patterns observed in the analyzed literature suggest that journalism scholars and media practitioners perceive deepfakes as a unique threat because they undermine visual and auditory evidence traditionally regarded as reliable indicators of reality. Unlike conventional fake news, which often relies on textual manipulation, deepfakes create highly convincing fabricated representations capable of mimicking authentic events and individuals. Such characteristics increase the potential for audience confusion and contribute to uncertainty regarding the credibility of digital information.

Thematic coding identified five major dimensions through which deepfakes influence epistemic trust in digital journalism. These dimensions include authenticity uncertainty, audience skepticism, verification challenges, reputational vulnerability, and institutional trust erosion. Authenticity uncertainty emerged as the most frequently discussed theme, appearing in 82% of the analyzed documents. Verification challenges followed at 77%, while audience skepticism was identified in 74% of the sources.

Institutional trust erosion appeared in 68% of the documents, indicating widespread concern regarding the long-term implications of synthetic media for journalism. Reputational vulnerability emerged in 63% of the sources and was frequently associated with political communication, public figures, and media organizations. Findings demonstrate that deepfakes generate multiple interconnected challenges that extend beyond individual instances of misinformation and affect broader processes of knowledge verification and public trust formation.

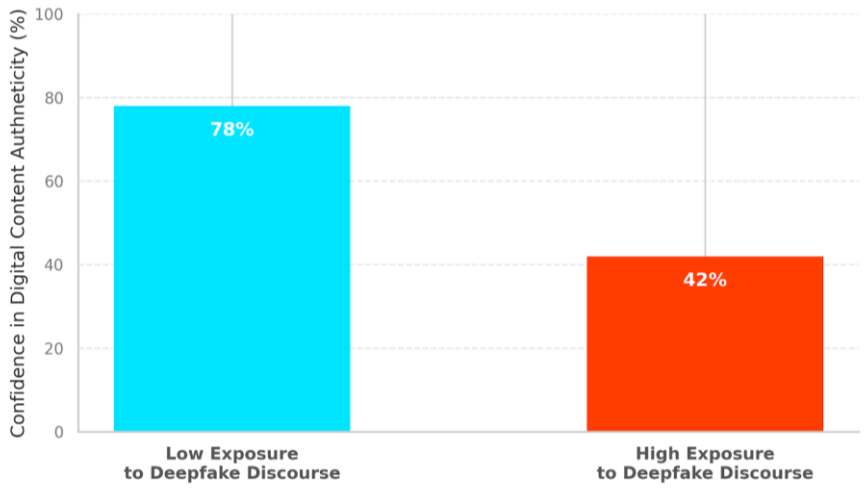


Figure 1. Impact of Deepfake Awareness on Digital Content Confidence

Interpretive analysis suggests a strong conceptual relationship between increased exposure to deepfake-related discourse and declining confidence in the authenticity of digital content. Studies reviewed consistently indicate that awareness of synthetic media capabilities influences audience evaluations of information credibility. Individuals familiar with the existence of deepfakes tend to exhibit greater skepticism toward audiovisual materials, even when presented with authentic journalistic content.

Analytical comparison across studies further suggests that epistemic trust is influenced not only by the presence of manipulated content but also by uncertainty regarding the possibility of manipulation. Findings reveal that suspicion alone may reduce confidence in information sources regardless of whether deception has actually occurred. This phenomenon contributes to what several scholars describe as a "liar's dividend," whereby genuine evidence can be dismissed as fabricated due to widespread awareness of deepfake technologies.

Relationships among the identified themes reveal a cyclical interaction between technological sophistication and trust erosion. Improvements in deepfake realism increase the difficulty of verification, which subsequently contributes to audience skepticism and uncertainty. Reduced confidence in digital content weakens trust in journalistic institutions, creating conditions in which misinformation and disinformation can spread more effectively.

Connections between audience skepticism and institutional credibility are particularly evident in the reviewed literature. Declining trust in media organizations often corresponds with increasing reliance on alternative information sources, some of which lack professional verification standards. Such dynamics reinforce the challenges faced by journalism in maintaining authority and legitimacy within increasingly fragmented information environments. Findings therefore suggest that technological developments and trust-related outcomes operate as mutually reinforcing processes.

Case study analysis focused on three widely discussed deepfake incidents involving political communication, public figures, and news dissemination. The first case involved manipulated video

content depicting a political leader delivering fabricated statements. The second case concerned a synthetic media campaign targeting a public figure during a period of heightened public attention. The third case examined a deepfake video that circulated widely across social media platforms before being debunked by professional fact-checkers and journalists.

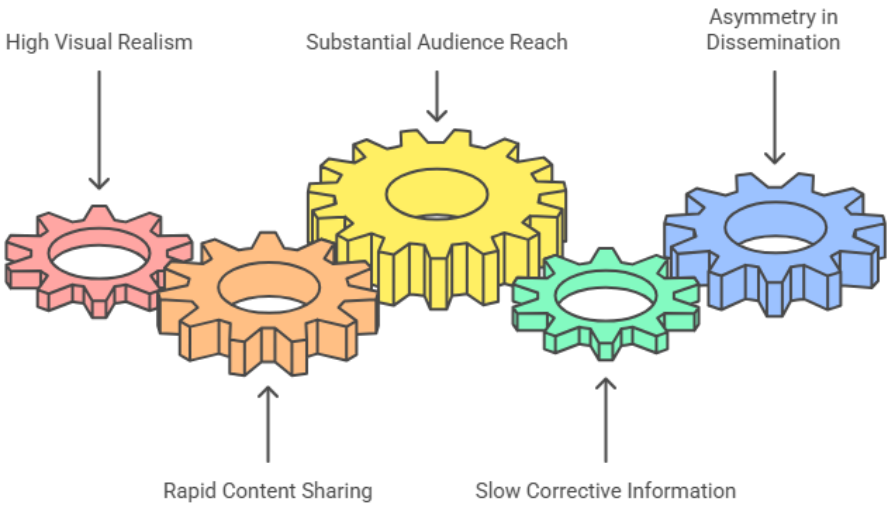


Figure 2. Misinformation Dissemination and Verification Process

Examination of these cases reveals similar patterns of dissemination and audience response. High levels of visual realism contributed to rapid content sharing before verification processes could be completed. Public engagement metrics demonstrated that manipulated content often reached substantial audiences within a short period, whereas corrective information generally circulated more slowly. These findings highlight the asymmetry between the speed of misinformation dissemination and the pace of journalistic verification.

Analysis of the selected case studies indicates that the effectiveness of deepfakes lies not only in their deceptive quality but also in their ability to exploit existing social and political tensions. Manipulated content frequently gains traction when it aligns with preexisting beliefs, ideological preferences, or emotional reactions among audiences. Such conditions increase the likelihood that deepfake content will be accepted, shared, and amplified before verification occurs.

Institutional responses observed across the cases demonstrate the growing importance of collaborative verification mechanisms involving journalists, technology platforms, fact-checking organizations, and academic researchers. Detection technologies alone were often insufficient to prevent the spread of manipulated content. Successful mitigation strategies typically combined technological tools with transparent communication, public education, and professional journalistic practices aimed at restoring confidence in verified information.

Overall findings indicate that deepfakes represent a transformative challenge for digital journalism because they undermine the epistemic foundations upon which public trust depends. Traditional assumptions regarding the reliability of audiovisual evidence are increasingly questioned within environments where synthetic media can closely imitate reality. Erosion of confidence in visual authenticity contributes to broader uncertainty regarding the credibility of information and institutions responsible for producing knowledge.

Interpretation of the results suggests that the crisis of epistemic trust extends beyond isolated incidents of misinformation. Deepfakes contribute to a structural transformation in how audiences evaluate evidence, credibility, and journalistic authority. Preservation of trust in digital journalism therefore requires not only improved detection technologies but also strengthened verification

practices, media literacy initiatives, ethical governance frameworks, and institutional transparency capable of supporting public confidence in an era of increasingly sophisticated synthetic media.

The findings reveal that deepfakes have evolved beyond the conventional boundaries of fake news and now represent a broader challenge to epistemic trust in digital journalism. Analysis of the selected literature demonstrates that concerns surrounding deepfakes are no longer limited to misinformation and disinformation but increasingly involve questions regarding authenticity, credibility, and the reliability of digital evidence. Deepfake technologies challenge one of the fundamental assumptions of journalism: the belief that visual and auditory records can function as trustworthy representations of reality.

The results indicate that authenticity uncertainty emerged as the most dominant concern across the analyzed studies. Audiences exposed to discussions about deepfakes often become more skeptical toward digital content regardless of whether the information is genuine or manipulated. Such skepticism reflects a transformation in how individuals evaluate information, particularly within environments characterized by rapid technological innovation and information overload.

The study further found that verification challenges constitute a significant obstacle for contemporary journalism. Traditional fact-checking mechanisms frequently struggle to keep pace with the speed and sophistication of synthetic media production. Journalists are increasingly required to verify not only textual claims but also highly realistic audiovisual materials that can closely imitate real events, public figures, and institutional communications.

Case study findings reinforce the broader thematic analysis by demonstrating that deepfakes can rapidly circulate through digital platforms before verification processes are completed. Patterns observed across multiple incidents indicate that manipulated content often receives greater initial attention than subsequent corrections. Such findings suggest that deepfakes possess the capacity to disrupt information ecosystems and weaken public confidence in journalistic institutions.

The findings align with earlier scholarship emphasizing the disruptive effects of misinformation on democratic communication and public trust. Previous studies have consistently argued that fake news undermines confidence in journalism, political institutions, and expert knowledge. Evidence generated through this study supports these conclusions while extending them to the context of synthetic media. Deepfakes amplify existing concerns by introducing highly convincing forms of audiovisual deception capable of bypassing traditional credibility assessments.

Research focusing on misinformation generally identifies false narratives, partisan content, and algorithmic amplification as primary drivers of trust erosion. Similar mechanisms are evident in the context of deepfakes. Manipulated videos and synthetic audio recordings frequently exploit emotional reactions, ideological biases, and existing social divisions to achieve rapid dissemination. Findings therefore reinforce the broader literature concerning the relationship between digital deception and declining trust in public information.

Differences emerge when comparing the present findings with studies that primarily examine the technical dimensions of deepfakes. Existing technological research often focuses on detection systems, machine learning models, and cybersecurity solutions. Findings from the current study suggest that technological detection alone cannot fully address the deeper crisis of epistemic trust generated by synthetic media. Questions concerning audience perception, institutional credibility, and information legitimacy remain central regardless of advances in detection capabilities.

The study also contributes a distinct perspective by framing deepfakes as an epistemological challenge rather than solely a technological or communicative problem. Earlier research frequently evaluates deepfakes in terms of content authenticity and misinformation risks. Findings presented here indicate that the broader issue involves uncertainty regarding the very possibility of knowing

whether digital evidence is genuine. Such a perspective expands existing discussions and highlights the need for interdisciplinary approaches integrating journalism, communication, technology, and epistemology.

The findings signify a profound transformation in the relationship between technology, knowledge, and public trust. Deepfakes challenge long-standing assumptions regarding the evidentiary value of visual and auditory materials. Photographs, videos, and audio recordings have historically served as persuasive forms of proof within journalism. Growing awareness of synthetic media technologies weakens confidence in these traditional indicators of authenticity.

The results also signify the emergence of a broader epistemic vulnerability within digital societies. Information consumers increasingly operate within environments where authentic and manipulated content coexist in ways that are difficult to distinguish. Such conditions encourage skepticism and uncertainty, potentially reducing confidence in reliable sources of information. Trust therefore becomes more fragile and dependent upon complex verification processes.

Patterns identified in the analysis further signify a shift from information abundance to authenticity scarcity. Digital environments provide unprecedented access to information while simultaneously generating greater uncertainty regarding the truthfulness of available content. Deepfakes exemplify this paradox by increasing the volume of persuasive information while decreasing confidence in its reliability. Journalism consequently faces new challenges in maintaining its role as a trusted source of public knowledge.

Broader interpretation suggests that the crisis of epistemic trust reflects deeper tensions between technological innovation and social institutions. Rapid advances in artificial intelligence have expanded the capacity to create synthetic realities, whereas institutional mechanisms for verification and accountability have evolved more slowly. Findings therefore indicate a growing need to reconsider how trust is established, maintained, and protected within increasingly digital communication environments.

The implications of these findings are substantial for journalism professionals and media organizations. Verification procedures must evolve to address increasingly sophisticated forms of synthetic media. Reliance on traditional fact-checking methods may no longer be sufficient in environments where manipulated audiovisual content can closely imitate authentic events. Investment in forensic verification tools and specialized training is therefore becoming increasingly important.

Educational implications are equally significant. Media literacy programs should move beyond conventional discussions of fake news and include deeper engagement with synthetic media technologies. Citizens require the skills necessary to critically evaluate digital content, understand the capabilities of artificial intelligence, and recognize the limitations of visual evidence. Strengthening public awareness may help mitigate some of the risks associated with deepfake dissemination.

Policy implications extend to regulatory frameworks governing artificial intelligence and digital communication. Governments and international organizations face increasing pressure to develop policies capable of balancing innovation with accountability. Regulatory approaches focused solely on content removal may prove insufficient because the broader challenge involves maintaining public trust rather than merely eliminating individual instances of deception.

Academic implications also emerge from the findings. Research concerning journalism, communication, and technology should increasingly incorporate epistemological perspectives to better understand how trust is shaped in digital environments. Future scholarship may benefit from examining the interactions among technological innovation, institutional legitimacy, audience

perception, and democratic participation. Such investigations could contribute to a more comprehensive understanding of contemporary information challenges.

The findings appear in this manner largely because deepfakes exploit cognitive tendencies that influence how individuals process information. Humans frequently rely on visual and auditory cues when assessing credibility. Synthetic media technologies capitalize on these tendencies by producing content that appears authentic even when entirely fabricated. Such characteristics make deepfakes particularly effective as tools of deception and manipulation.

Social media environments further contribute to the observed outcomes. Digital platforms prioritize speed, engagement, and virality, creating conditions that favor the rapid dissemination of emotionally compelling content. Verification processes often require time, expertise, and institutional resources. Manipulated content can therefore reach large audiences before corrective information becomes available. Structural features of digital communication thus amplify the influence of deepfakes.

Psychological factors also help explain the relationship between deepfakes and epistemic trust. Exposure to repeated examples of manipulated content encourages generalized skepticism toward digital information. Individuals may begin questioning not only suspicious materials but also authentic journalistic reporting. This phenomenon contributes to the erosion of confidence observed throughout the analyzed literature and case studies.

Institutional dynamics provide an additional explanation. Journalism developed within historical contexts where audiovisual evidence possessed relatively high credibility. Artificial intelligence technologies have altered these assumptions by enabling the creation of realistic synthetic content at unprecedented scale and speed. Institutional norms, professional practices, and regulatory frameworks have struggled to adapt at the same pace, creating the conditions necessary for an epistemic trust crisis to emerge.

Future efforts should prioritize the development of integrated verification ecosystems combining technological detection tools, journalistic expertise, platform accountability, and public education. Effective responses to deepfakes require collaboration among multiple stakeholders rather than reliance on any single solution. Comprehensive strategies are more likely to strengthen trust than isolated technological interventions.

Research should increasingly explore audience responses to synthetic media across different cultural, political, and informational contexts. Existing knowledge remains limited regarding how diverse populations interpret deepfakes and adapt their trust judgments over time. Longitudinal investigations could provide valuable insights into the evolving relationship between technological awareness and information credibility.

Media organizations should consider implementing transparent verification protocols capable of demonstrating how information is authenticated before publication. Greater openness regarding journalistic processes may help reinforce public confidence in professional reporting. Transparency initiatives could also contribute to rebuilding trust in institutions facing growing skepticism within digital environments.

Long-term resilience against the epistemic challenges posed by deepfakes will depend upon the ability of societies to cultivate informed, critical, and digitally literate citizens (Casás & Fernández Lombao, 2025). Strengthening trust cannot rely solely on technological safeguards because trust is ultimately a social and institutional phenomenon. Sustainable solutions therefore require coordinated efforts involving education, journalism, technology governance, and democratic institutions working collectively to preserve the integrity of public knowledge in the age of artificial intelligence.

CONCLUSION

The most important finding of this study is that deepfakes represent a fundamentally different challenge from conventional fake news because they do not merely spread false information; they destabilize the epistemic foundations upon which digital journalism depends. Findings indicate that the primary threat posed by deepfakes lies in their capacity to generate authenticity uncertainty, causing audiences to question not only manipulated content but also legitimate journalistic materials. This phenomenon contributes to a broader crisis of epistemic trust in which confidence in visual evidence, institutional verification, and journalistic authority becomes increasingly fragile. Deepfakes therefore function as catalysts for a structural transformation in how truth, credibility, and knowledge are evaluated within contemporary digital communication environments.

The principal contribution of this research is conceptual rather than purely methodological. Existing studies frequently examine deepfakes as technological artifacts, cybersecurity risks, or forms of misinformation. This study advances the discussion by positioning deepfakes within the framework of epistemic trust and digital journalism, thereby highlighting their broader implications for public knowledge, media legitimacy, and democratic discourse. Integration of perspectives from journalism studies, communication theory, artificial intelligence governance, and epistemology provides a more comprehensive understanding of how synthetic media reshapes trust relationships between information producers and audiences. Such a perspective contributes to the growing body of scholarship seeking to explain the societal consequences of generative artificial intelligence beyond technical considerations alone.

Several limitations should be acknowledged when interpreting the findings of this study. Reliance on secondary literature, policy reports, and documented case studies limits the ability to directly measure audience perceptions and behavioral responses to deepfakes. Rapid technological developments in generative artificial intelligence may also alter the nature of epistemic challenges faster than academic research can fully capture. Future investigations should incorporate empirical approaches involving audience surveys, experimental designs, cross-cultural comparisons, and longitudinal analyses to better understand how trust evolves in response to synthetic media exposure. Further research examining the effectiveness of media literacy programs, verification technologies, platform governance mechanisms, and journalistic interventions would provide valuable insights into strategies for strengthening epistemic resilience in increasingly complex digital information ecosystems.

DECLARATION OF AI AND AI ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this manuscript, the author(s) used QuillBot to assist inimproving grammar, language quality, and overall readability of the text. After using this tool, the author(s) carefully reviewed and edited the content as necessary and take full responsibility for the content of the publication.

AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.
Author 2: Conceptualization; Data curation; In-vestigation.
Author 3: Data curation; Investigation.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Abraham, T. M., Wen, T., Wu, T., & Chen, Y. (2025). Leveraging data analytics for detection and impact evaluation of fake news and deepfakes in social networks. *Humanities and Social Sciences Communications*, *12*(1), 1040. https://doi.org/10.1057/s41599-025-05389-4

Allen, J., Pennycook, G., & Rand, D. G. (2025). Addressing misperceptions takes more than combating fake news. *Trends in Cognitive Sciences*, *29*(9), 779–782. https://doi.org/10.1016/j.tics.2025.07.002

Anani-Bossman, A. (2026). Strategic Communication in the Era of Synthetic Media: Exploring Organizational Responses to AI-Generated Threats in Ghana. *International Journal of Strategic Communication*, 1–24. https://doi.org/10.1080/1553118X.2026.2614589

Avides Moreira, V., Hendrickx, J., & Schapals, A. K. (2026). "Talk to Me as a Friend!": How Teenagers Prefer Their Newsfluencers on Social Media. *Journalism and Media*, *7*(1), 19. https://doi.org/10.3390/journalmedia7010019

B, S., B, S., G S, S., Agarwal, H., & Bansal, P. (2025). TruthLens: A Serverless Deepfake Detection Platform Leveraging MesoNet and AWS Cloud Services. *2025 9th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, 1–5. https://doi.org/10.1109/CSITSS67709.2025.11294340

Bahra, M., & Fennan, A. (2025). Beyond Binary Sentiment: Leveraging Fuzzy Linguistic Hedges to Enhance NLP-Based Fake News Detection. Dalam B. E. Elbaghazaoui, M. Amnai, & N. Gherabi (Ed.), *Advances in Computational Intelligence and Robotics* (hlm. 27–42). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-8011-7.ch002

Casás, J., & Fernández Lombao, T. (2025). Estancamiento del fact-checking en los medios públicos desde la pandemia hasta hoy: Los casos de ZDF, FT y BBC. *Revista de Comunicación*, *24*(2), 57–82. https://doi.org/10.26441/RC24.2-2025-3992

Cazzamatta, R. (2025). Artificial Intelligence and Misinformation, Disinformation, Fake News, and Deepfakes. Dalam A. Sarısakaloğlu & M. Löffelholz (Ed.), *The Handbook of Artificial Intelligence and Journalism* (1 ed., hlm. 235–249). Wiley. https://doi.org/10.1002/9781394250424.ch16

Chen, X. K., & Na, J.-C. (2025). A Theory-Driven Approach to Fake News/Information Disorder Analysis and Explanation via Target-Based Emotion–Stance Analysis (TESA) and Interpretive Graph Generation (IGG). *Social Science Computer Review*, 08944393251338403. https://doi.org/10.1177/08944393251338403

Cheng, C. Y., & Hsu, C. H. (2025). Substitution or complementarity? Understanding the role of ChatGPT in transforming news media traffic in the United States and Taiwan. *Data Technologies and Applications*, *59*(4–5), 641–669. https://doi.org/10.1108/DTA-02-2025-0151

Díaz, F., Cerna, N., Liza, R., & Motta, B. (2026). Artificial Intelligence for Detecting Electoral Disinformation on Social Media: Models, Datasets, and Evaluation. *Information*, *17*(3), 292. https://doi.org/10.3390/info17030292

Dinana, H., Ali, D. A., & Taher, A. (2025). Trust Pathways in Digital Journalism: Comparing Western and National News Media Influence on Civic Engagement in Egypt. *Journalism and Media*, *6*(2), 61. https://doi.org/10.3390/journalmedia6020061

Doan, M. A. (2026). Issues, Risks, and Crisis Communication in the Age of 'Permacrisis.' Dalam G. Donnar & L. Parker, *Professional Communication* (1 ed., hlm. 92–103). Routledge. https://doi.org/10.4324/9781003543589-11

Dwyer, D. L. (2024). Watchdogs in the Digital Age: Digital Surveillance, Information Security, and the Evolution of Journalist-Confidential Source Relationships. Dalam M. Ambinder, J. R. Henrichsen, & C. Rosati (Ed.), *National Security, Journalism, and Law in an Age of Information Warfare* (1 ed., hlm. 125–154). Oxford University PressNew York. https://doi.org/10.1093/oso/9780197756621.003.0007

Fallis, D. (2025). Adversarial Epistemology. Dalam J. Lackey & A. McGlynn (Ed.), *The Oxford Handbook of Social Epistemology* (1 ed., hlm. 219–238). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190949945.013.11

Fuchs, C. (2025). What is and how do we achieve a resilient digital democracy? *Open Research Europe*, *5*, 387. https://doi.org/10.12688/openreseurope.21988.1

Hameed, H. Y., & Khudhair, B. W. (2026). The Ethical Challenges of Employing Artificial Intelligence in Digital Media from the Perspective of Media Professors in Iraq. Dalam A. Abdelgawad, A. A. Hameed, & A. Jamil (Ed.), *Intelligent Systems, Blockchain, and Communication Technologies* (Vol. 1801, hlm. 541–557). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-15784-3_40

Hoffmann, C. P., Bendahan Bitton, D., & Godulla, A. (2025). The Role of Self-Efficacy and Intellectual Humility in the Relationship Between Perceived Deepfake Exposure and Media Cynicism. *Social Science Computer Review*, 08944393251354977. https://doi.org/10.1177/08944393251354977

Hynek, N., Gavurova, B., & Kubak, M. (2025). Risks and benefits of artificial intelligence deepfakes: Systematic review and comparison of public attitudes in seven European Countries. *Journal of Innovation & Knowledge*, *10*(5), 100782. https://doi.org/10.1016/j.jik.2025.100782

Imran, H., Maqsood, M., & Imran, Q. (2025). The deepfake crisis: Exploring new media literacy, information integrity, and trust among Pakistani youth. *Qualitative Research Reports in Communication*, 1–12. https://doi.org/10.1080/17459435.2025.2541214

Ismail, A. I., Ismail, M. A. H., & Abu Hassan, N. F. (2025). The Mobilegraphy Era: Ethics, Innovation, and the Transformation of Visual Culture and Journalism. *PaperASIA*, *41*(4b), 243–253. https://doi.org/10.59953/paperasia.v41i4b.515

Karanja, J. M., & Nzau, T. (2026). The Emergence of AI Deepfakes and Their Impact on Strategic Crisis Communication: Practitioners' Perspectives From Kenya. *Journal of Contingencies and Crisis Management*, *34*(1), e70140. https://doi.org/10.1111/1468-5973.70140

Kassymbekova, N., Tolegen, E., Buyenbayeva, Z., Almanova, N., & Shyngyssova, N. (2025). The Impact of the Technological Boom on Traditional and Social Media in Kazakhstan. *Studies in Media and Communication*, *13*(2), 123. https://doi.org/10.11114/smc.v13i2.7536

Kim, H. V. T. (2026). The Use of AI in Newsrooms A Systematic Literature Review. Dalam P. Chien Thang & A. Prasanth (Ed.), *AI-Assisted Journalism and Media: Opportunities and Challenges* (Vol. 1277, hlm. 97–114). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-19648-4_7

McBrayer, J. P. (2026). *Misinformation: The Basics* (1 ed.). Routledge. https://doi.org/10.4324/9781003610779

Muralidharan, S. (2025). An Epidemiology of Fake News and a Possible Cure. Dalam P. Devi & A. Kaushik, *The Paradoxes of Free Speech* (1 ed., hlm. 75–91). Routledge. https://doi.org/10.4324/9781003541202-7

Murugalakshmi, V., Venkadesh, P., Divya, S. V., Gopinath, T., Jesso, P., & Raja, T. S. (2025). TruthChain: A Decentralized News Verification System Using Blockchain and AI. *2025 8th International Conference on Circuit, Power & Computing Technologies (ICCPCT)*, 1951–1956. https://doi.org/10.1109/ICCPCT65132.2025.11176464

Ramos-Zaga, F. (2025a). La crisis epistémica de la prueba audiovisual en la era de los deepfakes: Marco procesal-normativo para la autenticación y valoración judicial. *Revista Eletrônica de Direito Processual*, *27*(1). https://doi.org/10.12957/redp.2026.94995

Ramos-Zaga, F. (2025b). La crisis epistémica de la prueba audiovisual en la era de los deepfakes: Marco procesal-normativo para la autenticación y valoración judicial. *Revista Eletrônica de Direito Processual*, *27*(1). https://doi.org/10.12957/redp.2026.94995

Saranya, M. L., Haneesh, K. S., Kaushik, M., Vanitha, M. L., Srinivas, A., & SivaKrishna, K. (2026). Veritas: Real-Time News Summarization and Credibility Analysis Using Machine Learning. *2026 International Conference on Machine Learning and Autonomous Systems (ICMLAS)*, 172–176. https://doi.org/10.1109/ICMLAS67792.2026.11483964

Teixeira, M. (2025). TECHNOLOGIES, DISINFORMATION AND KAFKA IN THE DIGITAL AGE. *e3*, *10*(1), 95–109. https://doi.org/10.29073/e3.v11i1.1005

Valenzuela, S., Bachmann, I., Borah, P., & Solís Valdés, N. (2026). The Effects of Generative AI in News on Media Credibility and Selectivity: Evidence from a Conjoint Experiment in Chile. *Digital Journalism*, 1–22. https://doi.org/10.1080/21670811.2026.2649018

Vineela, A., Bhavani, A., Krishna, B. V., & Sankar, A. B. (2025). An artful multimodal exploration in discerning fake news through text and image harmony. *Multimedia Tools and Applications*, *84*(31), 37997–38016. https://doi.org/10.1007/s11042-025-20695-4

Wanjari, U., Kadam, S. K., Raj, C., Anawade, P., & Sharma, D. (2026). Technological and Policy Approaches to Combatting Disinformation Online. Dalam J. Choudrie, P. N. Mahalle, T. Perumal, & A. Joshi (Ed.), *ICT for Intelligent Systems* (Vol. 1507, hlm. 13–24). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-8750-3_2

Whitelaw, B. (2026). The Three Eras of Content Moderation in the Media and What Comes Next. Dalam M. L. Daniel, A. Menking, M. Savio, & J. Claffey, *Trust, Safety, and the Internet We Share* (1 ed., hlm. 117–125). Routledge. https://doi.org/10.1201/9781003621072-15

Wirtz, B. W., Weyerer, J. C., & Müller, T. F. (2026). AI-based digital disinformation: A theory-informed and integrated trilateral framework of digital disinformation for information systems research. *International Journal of Information Management*, *89*, 103066. https://doi.org/10.1016/j.ijinfomgt.2026.103066

Yang, B. (2025). Unveiling the Potential of AI-Driven Social Media: Applications of Intelligent Data Analysis in News Communication. *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, 1–6. https://doi.org/10.1109/ICMSCI62561.2025.10894172

Yazdi, S. V., Gong, Y., & Chen, Q. (2026). Anger or love? Emotions as complementary signals beyond text for fake news detection. *Annals of Operations Research*. https://doi.org/10.1007/s10479-026-07179-w

Zhang, Y., Lin, W., Xu, J., Xu, W., & Xu, Y. (2025). MLPN: Multi-Scale Laplacian Pyramid Network for deepfake detection and localization. *Journal of Information Security and Applications*, *89*, 103965. https://doi.org/10.1016/j.jisa.2025.103965

Copyright Holder :
© Prasojo et al. (2026).

First Publication Right :
© Journal of Social Science Utilizing Technology

This article is under:

