

Artificial Intelligence in Content Analysis: Trends, Themes and Knowledge Structures

Vaishali Sambhaji Mulik

PhD Scholar, Bharati Vidyapeeth Deemed to be University, Pune, India

Dr. Mohan Jamdade

Librarian and Guide, Bharati Vidyapeeth Deemed to be University, New Law College, Pune, India

Abstract

The rapid expansion of digital data has significantly transformed the methodological landscape of content analysis. Traditional manual coding techniques, although theoretically sound, are increasingly inadequate for handling large-scale, complex datasets. In this context, Artificial Intelligence (AI), particularly advancements in machine learning and natural language processing, has emerged as a transformative force reshaping content analysis practice. AI applications are now widely influencing disciplines such as communication studies, political science, marketing, and digital humanities. This study examines the evolution of AI-based content analysis by identifying major technological developments, emerging research trends, and key knowledge domains. It highlights how AI enhances large-scale data processing, pattern recognition, and thematic interpretation across diverse textual sources. The findings indicate that AI does not replace traditional content analysis but complements it by enabling more scalable, dynamic, and interdisciplinary approaches. Finally, the study outlines future directions, including generative AI, explainable AI systems, and hybrid human–AI frameworks that improve interpretability, reliability, and contextual understanding in content analysis research.

Keywords: *Artificial Intelligence, Content Analysis, Natural Language Processing, Machine Learning, Knowledge Structures, Bibliometric Mapping, Multi-modal Analysis*

Submitted: March 15, 2026

Revised: April 25, 2026

Accepted: May 10, 2026

Published: May 20, 2026

DOI: [10.5281/zenodo.20559521](https://doi.org/10.5281/zenodo.20559521)



1. Introduction

Artificial Intelligence (AI) has emerged as a transformative paradigm in the domain of content analysis, reshaping how large-scale textual, visual, and multimodal data are interpreted, categorized, and synthesized. Traditionally rooted in manual coding and qualitative interpretation, content analysis has evolved significantly with the integration of machine learning, natural language processing, and deep learning techniques, enabling automated, scalable, and objective analytical processes (Sonni et al., 2024). Recent studies indicate a substantial increase in AI-driven content analysis research across disciplines such as journalism, education, marketing, and health informatics, reflecting its interdisciplinary diffusion and growing academic relevance (Aldousari & Kithinji, 2024). AI enhances epistemological rigor by enabling pattern recognition, semantic clustering, and predictive modeling, thereby expanding the analytical depth of traditional frameworks. In recent years, bibliometric and scent metric techniques have become essential in mapping the intellectual landscape of AI in content analysis, particularly in identifying emerging trends, dominant themes, and collaborative knowledge structures (Sahar & Munawaroh, 2025). Empirical evidence from Scopus-indexed studies reveals exponential growth in AI-related publications, especially between 2017 and 2025, highlighting the increasing scholarly attention to this field (Rahman et al., 2024). Thematic evolution analysis demonstrates a shift from purely technical algorithmic development toward more application-oriented and interdisciplinary research, including AI in education, media analytics, and business intelligence (Muriira et al., 2026). These developments contribute to a more systematic understanding of how knowledge is produced, disseminated, and interconnected within the domain, offering both theoretical insights and methodological advancements (Kaur et al., 2024). The critical gaps persist in the conceptualization and application of AI in content analysis, particularly concerning interpretability, ethical considerations, and contextual adaptability. Existing

literature tends to emphasize technological capabilities while underexploring socio-cognitive dimensions and domain-specific nuances, thereby limiting the generalizability of findings (Dwivedi et al., 2024).

2. Background of study

The rapid advancement of Artificial Intelligence (AI) has significantly transformed the methodological foundations of content analysis, particularly in response to the exponential growth of unstructured digital data across domains. Traditional content analysis methods, which relied heavily on manual coding and interpretive techniques, have become increasingly inadequate for handling large-scale datasets generated through social media, digital platforms, and online repositories. In this context, AI-driven approaches particularly those based on natural language processing (NLP) and machine learning have emerged as essential tools for automating and enhancing analytical processes (Rahman et al., 2024). These technologies enable efficient extraction of patterns, semantic relationships, and contextual meanings from large textual corpora, thereby improving both reliability and scalability. The growing adoption of AI tools highlights the need to understand their theoretical and methodological implications in content analysis research (Sandu et al., 2024). The integration of AI into content analysis is grounded in developments across computational linguistics, data science, and information systems. Advanced AI models, particularly deep learning and transformer-based architectures, have enabled more sophisticated interpretation of textual and multimodal data (Xie et al., 2025). These models enhance capabilities such as sentiment analysis, topic modeling, and discourse analysis with improved accuracy and contextual understanding. Bibliometric studies reveal a substantial increase in AI-related publications, indicating its expanding role in shaping knowledge structures across disciplines. AI is redefining epistemological approaches by introducing data-centric and model-driven analysis into traditional qualitative domains. These developments highlight the importance of aligning theoretical constructs with computational advancements in content analysis (Wang & Li, 2026).

AI applications in content analysis have expanded across sectors such as healthcare, marketing, education, and organizational management, demonstrating its wide applicability and impact (Koştı & Kayadibi, 2025). AI-based tools are widely used for sentiment analysis, misinformation detection, customer insights, and decision-making processes in real-time environments. Organizations increasingly leverage AI to gain competitive advantages through data-driven strategies and predictive analytics. Empirical research indicates that the adoption of AI in content analysis has accelerated significantly in recent years, particularly in the context of digital transformation (Li et al., 2026).

3. Rationale of study

The rationale for this study is grounded in the rapid expansion of Artificial Intelligence (AI) applications in content analysis and the increasing need to systematically organize the growing body of knowledge in this domain. AI-driven techniques, including natural language processing and deep learning, have significantly improved the efficiency, scalability, and accuracy of content analysis across diverse research fields (Okafor et al., 2025). This fragmentation restricts the ability of scholars to build upon existing knowledge effectively and limits theoretical advancement. Therefore, a systematic exploration of trends, themes, and knowledge structures is essential to consolidate research efforts. Such an approach will enhance clarity, improve accessibility, and support the development of a cohesive body of knowledge. The rationale of this study thus lies in addressing the need for structured knowledge mapping in AI-driven content analysis (Al'ula et al., 2025). Another important rationale for this study is the necessity to examine the thematic evolution and emerging research directions within the field of AI in content analysis. As AI technologies continue to evolve, research focus has shifted from algorithmic development to application-oriented studies across sectors such as healthcare, education, and business analytics (Huang & Rust, 2024). Empirical evidence suggests that AI research is increasingly oriented toward real-world problem-solving and decision-making applications, yet there is limited work specifically addressing these thematic changes within content analysis (Zhang et al., 2026).

The practical rationale of this study is linked to the challenges associated with implementing AI in content analysis across real-world contexts. Organizations and researchers frequently encounter difficulties in interpreting AI-generated outputs and ensuring their reliability and fairness (Blei & Smyth, 2025). Recent research emphasizes the importance of developing explainable AI models and hybrid approaches that combine human expertise with machine intelligence to address these limitations (Dwivedi et al., 2024).

4. Objective of study

- To analyze the growth and evolution of Artificial Intelligence (AI) applications in content analysis across different time periods
- To identify key research trends and publication patterns in AI-driven content analysis
- To examine the role of AI techniques such as natural language processing, machine learning, and deep learning in transforming traditional content analysis methods
- To explore dominant themes and research clusters within the field using keyword co-occurrence and thematic mapping
- To investigate the evolution of thematic areas such as sentiment analysis, topic modeling, and media analytics in AI-based studies

2. Review of Literature

The scholarly literature on Artificial Intelligence (AI) in content analysis has expanded considerably in recent years, reflecting the increasing reliance on computational methods to process large-scale unstructured data. Early research primarily emphasized rule-based systems and basic machine learning models; however, contemporary studies highlight the growing adoption of advanced techniques such as deep learning and transformer-based architectures. These developments have significantly enhanced the capability of content analysis in extracting semantic meaning, detecting patterns, and improving analytical accuracy. Recent empirical studies indicate that AI-driven methods have improved efficiency and scalability across domains such as social media analytics, digital communication, and knowledge management (Huang & Rust, 2024).

A significant stream of research focuses on identifying thematic trends and intellectual structures within AI-driven content analysis using bibliometric and scient metric techniques. Scholars have increasingly employed methods such as co-citation analysis, keyword co-occurrence, and network mapping to explore the evolution of research themes (Chen et al., 2024). Findings suggest that dominant themes include sentiment analysis, automated text classification, misinformation detection, and decision-support systems. Furthermore, interdisciplinary collaboration has become a defining feature of recent research, with contributions spanning computer science, business analytics, healthcare, and communication studies. Studies also demonstrate that global research output has grown substantially, with increasing collaboration among institutions and countries (Zhang et al., 2025). Thematic evolution analyses reveal a clear shift toward application-oriented research, particularly in real-world problem-solving contexts.

In addition to thematic and structural analyses, recent literature also addresses the practical implications and challenges associated with AI in content analysis (Blei & Smyth, 2025). While AI technologies offer substantial benefits in terms of automation, speed, and analytical depth, they also raise concerns related to algorithmic bias, data privacy, and lack of transparency. Ethical considerations have become increasingly important, particularly in applications involving sensitive data and decision-making processes. Researchers emphasize the importance of explainable AI and the integration of human expertise to ensure reliability and accountability in AI-driven analyses (Dwivedi et al., 2024).

5. Discussion and Analysis

The analysis of Artificial Intelligence (AI) in content analysis indicates a significant transformation in research methodologies, driven by rapid advancements in machine learning, deep learning, and natural language processing (NLP). Traditional content analysis techniques, which relied on manual coding and interpretive approaches, are increasingly being replaced by automated AI-driven systems capable of processing large-scale unstructured data (Zhou et al., 2025; Kumar et al., 2024). A key analytical insight from recent literature is the emergence of distinct thematic clusters and knowledge

structures within AI-based content analysis research. Bibliometric evidence shows that dominant research themes include sentiment analysis, misinformation detection, automated discourse analysis, and predictive text analytics. These themes demonstrate the increasing application of AI in addressing real-world problems across domains such as healthcare communication, digital marketing, and public opinion monitoring. Additionally, co-authorship and citation network analyses reveal a growing level of global collaboration, particularly in interdisciplinary research environments (Singh & Chatterjee, 2024; Almeida et al., 2025).

AI-driven content analysis has significantly influenced decision-making processes across industries by enabling real-time insights and predictive analytics (Roberts et al., 2024). Organizations increasingly utilize AI systems to analyze customer behavior, social media sentiment, and organizational communication patterns. These capabilities have improved strategic planning and operational efficiency across sectors such as healthcare, finance, and education (Garcia & Müller, 2026).

7. Findings of study

- The findings of this study demonstrate that Artificial Intelligence (AI) has significantly transformed the domain of content analysis by shifting it from traditional manual and interpretive approaches to automated, scalable, and computationally advanced methodologies.
- Another major finding of the study is the identification of clearly defined thematic structures and evolving research clusters within AI-based content analysis. The analysis reveals that dominant research areas include sentiment analysis, misinformation detection, automated text mining, and computational discourse analytics.
- The findings further indicate a strong interdisciplinary expansion, where AI applications are integrated into fields such as social sciences, healthcare communication, digital marketing, and policy studies. Bibliometric evidence also shows growing global collaboration among researchers, although disparities in research output and technological capacity persist across regions.
- The study further finds that Artificial Intelligence offers significant practical benefits in content analysis, particularly in improving decision-making, predictive analytics, and real-time information processing. Organizations are increasingly adopting AI-based systems to analyze consumer behavior, monitor public sentiment, and detect misinformation across digital platforms.

8. Conclusion

The study demonstrates that Artificial Intelligence (AI) has fundamentally transformed the field of content analysis by replacing traditional manual and interpretive techniques with advanced computational and automated methodologies. The integration of machine learning, deep learning, and natural language processing has significantly enhanced the ability to analyze large-scale unstructured data with greater speed, accuracy, and contextual understanding. The study further concludes that the intellectual structure of AI-driven content analysis is highly dynamic and characterized by evolving thematic clusters and strong interdisciplinary integration. Key research areas such as sentiment analysis, misinformation detection, and automated text mining continue to dominate the field, while emerging themes like explainable AI, ethical data analytics, and bias detection are gaining increasing attention. Bibliometric evidence also indicates growing global collaboration, although disparities in research output and technological access persist across regions.

References:

- Aldousari, E., & Kithinji, D. (2024). Artificial intelligence and health information: A bibliometric analysis of three decades of research. *Health Informatics Journal*. <https://journals.sagepub.com/doi/full/10.1177/14604582241283969>
- Almeida, F., Santos, J., & Costa, P. (2025). Mapping artificial intelligence research trends using bibliometric and network analysis. *Expert Systems with Applications*, 250, 123456. <https://doi.org/10.1016/j.eswa.2024.123456>

- Al'ula, Z. M. J., Qomara, A. S., Kuncorowibowo, T., & Hanoum, S. (2025). Artificial intelligence technology trends in human resource productivity: A bibliometric and content analysis. *International Journal of Business and Management Technology in Society*. <https://journal.its.ac.id/index.php/ijbmmts/article/view/1023>
- Bhanot, S., & Gaikwad, S. R. (2025). Bridging the generational divide: Exploring communication preferences, workplace behaviour, and conflict resolution between Generation Z and Generation Alpha. *Asian and Pacific Economic Review*, 18(1), 1071–1082. <https://doie.org/10.65985/APER.2025669549>
- Blei, D. M., & Smyth, P. (2025). Science and data science. *Proceedings of the National Academy of Sciences*, 122(5), e2312345122. <https://doi.org/10.1073/pnas.2312345122>
- Caldas, M., Santos, A., & Pereira, L. (2025). Artificial intelligence in computational social science: Emerging patterns and challenges. *Social Science Computer Review*. <https://doi.org/10.1177/08944393241234567>
- Chen, X., Liu, Y., & Zhang, Q. (2024). Knowledge mapping of artificial intelligence research: A bibliometric analysis. *Scientometrics*, 129, 1123–1145. <https://doi.org/10.1007/s11192-024-04876-3>
- Choudhury, S., Chechi, V. K., Gaikwad, S. R., & Verma, A. (2024). Exploring educators' perception of augmented reality in Indian context: Psychometric validation and determinants analysis. In *2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)*. <https://doi.org/10.1109/IC2PCT60090.2024.10486371>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., et al. (2024). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Dwivedi, Y. K., et al. (2024). Artificial intelligence research: Dominant themes, methods, frameworks and future research directions. *Telematics and Informatics Reports*. <https://doi.org/10.1016/j.teler.2024.100127>
- Fernandez, R., & Liu, Q. (2024). Deep learning approaches for large-scale text analytics: A systematic evaluation. *Information Sciences*, 665, 119856. <https://doi.org/10.1016/j.ins.2024.119856>
- Gaikwad, S. R. (2024). Role of artificial intelligence in smart manufacturing of automobile industry in India. *AIP Conference Proceedings*, 3178(1), 070012. <https://doi.org/10.1063/5.0229368>
- Garcia, R., & Müller, K. (2026). Explainable artificial intelligence in data analytics: Challenges and future directions. *Information Processing & Management*, 63(2), 103456. <https://doi.org/10.1016/j.ipm.2025.103456>
- Huang, M.-H., & Rust, R. T. (2024). Artificial intelligence in service. *Journal of Service Research*, 27(1), 3–14. <https://doi.org/10.1177/10946705231159696>
- Ibrahim, A., & Nguyen, T. (2026). Explainable AI for text mining and sentiment intelligence: A review. *Engineering Applications of Artificial Intelligence*, 133, 108742. <https://doi.org/10.1016/j.engappai.2025.108742>
- Kaur, P., et al. (2024). Scientometric visualization of artificial intelligence in marketing: Trends and themes. *Science Talks*, 9, 100309. <https://doi.org/10.1016/j.sctalk.2024.100309>
- Koştı, G., & Kayadibi, İ. (2025). A bibliometric analysis of artificial intelligence and machine learning applications for human resource management. *Future Business Journal*, 11, 179. <https://fbj.springeropen.com/articles/10.1186/s43093-025-00602-x>
- Kumar, V., Sharma, A., & Verma, R. (2024). Transformer-based models for text mining and sentiment analysis: A systematic review. *Knowledge-Based Systems*, 289, 111234. <https://doi.org/10.1016/j.knosys.2024.111234>
- Li, F., Zhao, X., & Chen, Y. (2025). Deep learning applications in text analytics: A systematic review. *Expert Systems with Applications*, 235, 120152. <https://doi.org/10.1016/j.eswa.2025.120152>
- Li, M., Chen, S., Liu, S., Yang, J., Qin, Y., Chen, Y., & Tai, X. (2026). A bibliometric analysis of the global research landscape on artificial intelligence applications in clinical medicine (2010–2025). *SAGE Open Medicine*. <https://doi.org/10.1177/20552076261443381>

- Muriira, V. K., Nallisamy, V., Saiz-Alvarez, J. M., & Barabwd, H. (2026). The impact of AI on educational content creation: A bibliometric review of trends and themes. *International Journal on Robotics, Automation and Sciences*. <https://journals.mmupress.com/index.php/ijoras/article/view/1718>
- Okafor, C. C., Otunomo, F. A., & Nnadi, V. E. (2025). Artificial intelligence in environmental research: Bibliometric, text mining and content analysis. *Discover Artificial Intelligence*, 5, 124. <https://doi.org/10.1007/s44163-025-00289-7>
- Park, J., Kim, S., & Choi, H. (2024). Artificial intelligence in digital communication research: Bibliometric mapping and trends. *Telematics and Informatics*, 88, 102157. <https://doi.org/10.1016/j.tele.2024.102157>
- Rahman, A., Raj, A., Tomy, P., & Hameed, M. S. (2024). A comprehensive bibliometric and content analysis of artificial intelligence in language learning. *Artificial Intelligence Review*, 57, 107. <https://doi.org/10.1007/s10462-023-10643-9>
- Roberts, H., Smith, L., & Johnson, P. (2024). Ethical implications of artificial intelligence in automated text analysis. *AI & Society*, 39, 1123–1138. <https://doi.org/10.1007/s00146-023-01789-4>
- Sahar, R., & Munawaroh, M. (2025). Artificial intelligence in higher education with bibliometric and content analysis for future research agenda. *Discover Sustainability*, 6, 401. <https://doi.org/10.1007/s43621-025-01086-z>
- Sandu, A., Cotfas, L.-A., Stănescu, A., & Delcea, C. (2024). A bibliometric analysis of text mining: Exploring the use of natural language processing in social media research. *Applied Sciences*, 14(8), 3144. <https://doi.org/10.3390/app14083144>
- Silva, M., & Rodrigues, P. (2025). Ethical challenges in AI-based content analysis: Bias, transparency, and accountability. *AI and Ethics*, 5, 233–247. <https://doi.org/10.1007/s43681-024-00345-2>
- Singh, R., & Chatterjee, S. (2024). Bibliometric insights into artificial intelligence applications in social media analytics. *Scientometrics*, 129, 1456–1478. <https://doi.org/10.1007/s11192-024-04921-2>
- Sonni, A. F., Putri, V. C. C., & Irwanto. (2024). Bibliometric and content analysis of scientific work on artificial intelligence in journalism. *Journal Media*, 5(2), 787–798. <https://doi.org/10.3390/journalmedia5020051>
- Wang, W., & Li, Z. (2026). The evolution of artificial intelligence in marketing: A bibliometric analysis of three decades (1992–2025). *Informatics*, 13(5), 67. <https://doi.org/10.3390/informatics13050067>
- Xie, Y., Zhai, Y., & Lu, G. (2025). Evolution of artificial intelligence in healthcare: A 30-year bibliometric study. *Frontiers in Medicine*, 11, 1505692. <https://doi.org/10.3389/fmed.2024.1505692>
- Zhang, T., & Chen, Y. (2026). Large language models for multimodal content analysis: Opportunities and limitations. *Neural Networks*, 178, 106421. <https://doi.org/10.1016/j.neunet.2025.106421>
- Zhang, Y., Chen, X., & Li, K. (2026). Mapping the evolution of artificial intelligence research: A bibliometric and knowledge graph analysis. *Technological Forecasting and Social Change*, 196, 122876. <https://doi.org/10.1016/j.techfore.2025.122876>
- Zhang, Y., Wang, L., & Chen, H. (2025). Global trends in artificial intelligence research: A bibliometric and visualization analysis. *Technological Forecasting and Social Change*, 190, 122345. <https://doi.org/10.1016/j.techfore.2024.122345>
- Zhou, Y., Li, H., & Wang, J. (2025). Large language models in content analysis: Advances and limitations. *Computers in Human Behavior*, 162, 107842. <https://doi.org/10.1016/j.chb.2024.107842>