

Comparative Study on the Credibility of AI-Generated News Content and Traditional News Content

Aiyun Hua

School of Public Administration, Nanjing Normal University, Nanjing 210023, China

**Corresponding author: Aiyun Hua, E-mail: wxhaytt@163.com.*

Abstract

Currently, the credibility of AI-generated news content is one of the greatest problems in the development of AIGC technology in journalism. To explore this problem, this article compares the credibility of AI-generated news content and traditional news content in three main aspects: the amount and quality of used data, the personal bias of journalists, and the information transparency of the news content. Through case analysis, data citation, and critical thinking, this study reveals that AI-generated news benefits from vast training data and objective algorithms, reducing personal bias and enabling broad coverage. However, its credibility is undermined by the potential inclusion of inaccurate or biased data and a lack of transparency regarding algorithms and data sources. By examining both the advantages and disadvantages of AI-generated and traditional news content, this study provides ideas and thoughts for the future development of AIGC technology in journalism.

Keywords

AI-generated news content, traditional news content, credibility comparative study

1. Introduction

In recent years, artificial intelligence (AI), the simulation of human intelligence in machines or computer systems, has become one of the most popular topics and research objects worldwide. The arrival of the AI era is as significant as the first stone tool created by humans tens of thousands of years ago, meaning that people can start a brand-new work and life with the help of the new tool. The use of AI greatly reduces the investment of manpower and material resources. In fact, AI has not penetrated various industries, including journalism, transportation, education, entertainment, healthcare, etc.

In journalism, AI-generated news content is increasingly adopted by major media organizations, gradually replacing traditional news content. For instance, the Associated Press has used AI to generate financial reports since 2014 (Miller, 2015), and Bloomberg's "Cyborg" system has assisted journalists in producing business news. China's Xinhua News Agency introduced an AI-powered virtual news anchor capable of delivering news 24/7. While these advancements have improved efficiency, the credibility of AI-generated news is still doubted by people. There is high risk for artificial intelligence-generated content

(AIGC) technology to cause AI information pollution on various social media platforms, such as Facebook, Twitter, WeChat, and Weibo, which may greatly mislead the public (Han et al., 2024). This study aims to compare the credibility of AI-generated news content and traditional news content. I explore this issue from three aspects:

- 1) Amount and quality of data used.
- 2) Personal biases of journalists.
- 3) Information transparency of news content.

2. Amount and Quality of Used Data

2.1 Amount of Data

AI-generated news content relies on algorithms trained on vast datasets of text to autonomously create articles, which means that the data that it is based on are much larger than those of traditional news content. Using big data for news content offers several benefits, which is conducive for credibility. First, AI enhances accuracy by analysing extensive sources, ensuring comprehensive coverage of events. For example, AI-powered journalism platforms such as Bloomberg's "Cyborg" analyse financial data in real time, allowing for rapid and precise reporting. Second, AI can detect trends and gauge public sentiment by processing massive amounts of information, improving the relevance and depth of news reporting. Moreover, AI-generated news benefits from real-time analytics, enabling faster updates on breaking news, unlike traditional journalism, where verification and editorial processes often delay reporting.

2.2 Quality of Data

However, the quality of the data used by AIGC tends to be a concern. The database inevitably contains a certain proportion of erroneous data and biased information. For example, the accuracy of IP geolocation databases is often called into question. Different IP geolocation databases may yield inconsistent results for the same IP address (Xie et al., 2024). Johnson et al. (2023) reviewed 43 studies and reported that EHRs frequently presented incomplete and/or inaccurate data on the race/ethnicity of patients. AI-generated news content based on inaccurate or erroneous data is undoubtedly untrustworthy.

In addition, biased information contained in the database is also a major problem. AIGC technology is trained on archival data produced by humans, which means that it can inherit and even amplify biases presented in the training data. In the study of Fang et al. (2024), the authors investigated the bias of AIGC produced by seven representative large language models (LLMs) by comparing the AIGC and the original news articles from The New York Times and Reuters. They find that the AIGC produced by each examined LLM demonstrates substantial gender and racial biases due to biased information contained in training data, exhibiting notable discrimination against females and individuals of the Black race.

2.3 Comparison with Traditional News Content

In contrast, traditional news content benefits from professional journalists who verify information before publication, reducing the chances of spreading false or biased reports. For example, in 2020, *the New York Times* conducted an in-depth investigation into the origins of COVID-19, relying on expert interviews, government reports, and scientific studies to ensure accuracy. In contrast, AI-generated reports on the same topic, such as those produced by automated content systems on social media, occasionally spread misleading claims due to a reliance on unverified data (Nyilasy, 2020).

Another case is election coverage. In the 2024 U.S. presidential election, *Washington Post* assigned journalists to fact-check political claims, ensuring accuracy and context. Moreover, AI-generated news summaries on platforms such as Google News sometimes misrepresent pooled data by failing to account for margin-of-error adjustments, leading to misleading narratives.

These examples highlight the key difference: while AI-generated news benefits from vast data sources and real-time analytics, its reliance on potentially flawed databases poses a serious challenge. Despite its limitations, traditional journalism remains more reliable in ensuring data accuracy, objectivity, and credibility.

3. Personal Bias of Journalists

Personal bias is one of the most important factors influencing the credibility of news content. Everyone is an independent and unique individual. People have different genes, experiences, and backgrounds, all of which determine that everyone has varying levels and types of stereotypes and even biases. Traditional news content is completed by journalists, which means that the issue of personal bias can never be avoided. The coverage of the Trayvon Martin case (2012) is a good example. During the trial of George Zimmerman for the shooting death of Trayvon Martin, media coverage varied significantly. Some outlets framed the incident as a racially motivated killing, highlighting systemic racism and injustice in law enforcement. Others emphasized self-defense laws, portraying Zimmerman as someone exercising his legal rights. These different portrayals influenced how the public understood the case. Audiences exposed to racially charged narratives were more likely to view Martin as a victim of racial profiling, whereas those following self-defense narratives viewed Zimmerman's actions as justified. This division in coverage deepened social and political tensions, demonstrating how media framing can shape public discourse and societal debates.

Another example is immigration coverage, where journalistic bias influences the perceptions of migrants. Some media outlets focus on crime rates or economic burdens, reinforcing negative stereotypes and shaping public fears about immigration (Agovino et al., 2022). Others highlight humanitarian aspects, emphasizing the struggles and contributions of migrants. The way these narratives are framed directly impacts public opinion. If immigration is predominantly linked to crime, audiences may develop hostile attitudes and support restrictive policies. Conversely, if the narrative centers on human rights and economic benefits, the public may be more supportive of inclusive policies.

Compared with traditional news content, AI-generated news content does not have the serious problem of personal bias, which is positive for its credibility. Like what I have mentioned before, AIGC technology utilizes artificial intelligence algorithms trained on extensive datasets. The algorithms used here tend to be natural language processing (NLP) models. NLP is a branch of artificial intelligence that focuses on the interaction between computers and human (natural) languages and involves the application of computational techniques to analyse and model natural language data. Key tasks in the NLP model include language translation, sentiment analysis, text summarization, named entity recognition, and language generation (Al-Khalifa et al., 2024). The entire generation process of AI-generated news content, each task of AIGC technology, is not influenced by human emotions, desires, or personal biases, ensuring its relative objectivity and credibility in this respect. In other words, biased AI-generated news content can be caused only by biased information in the database and is unrelated to the generation process.

4. Information Transparency of News Content

Various countries have specific laws and commissions aimed at ensuring the information transparency of journalism. For example, in Germany, media transparency is regulated by the Press Council (Presserat), which issues guidelines on journalistic standards, including transparency in reporting and editorial independence. In the United Kingdom, the Communications Act of 2003 and guidelines from regulatory bodies such as Ofcom play important roles, ensuring that news media operations are accountable and transparent in their reporting. Australia has laws such as the Broadcasting Services Act 1992 and regulatory oversight by bodies such as the Australian Communications and Media Authority (ACMA). As a result, the information transparency of traditional news content is always protected comprehensively and legally, thereby building a secure bridge of trust with the audience.

However, all these laws, regulations and commissions can ensure only the information transparency of traditional news content, as AI generates news content in a completely different way. From the perspective of AIGC, "among the sorts of information targeted by the notion of information transparency is information

about the nature of the algorithms that are used by a given AI system, as well as the kinds of data that were used in a training regimen” (Andrada et al., 2023, p. 1324). Unfortunately, revealing the information of the algorithms used and training data may be a formidable technical challenge. Many computing systems programmed via machine learning (ML) are opaque, which means that it is difficult to “look inside” to explain their behavior. Even the designers of systems may not be certain exact methods by which certain system outputs are generated (Zednik, 2021). As a result, most countries have not implemented specific regulations on the information transparency of AIGC. The lack of information transparency directly reduces the credibility of AI-generated news content, which is an issue worth exploring for the future development of AIGC technology.

The lack of transparency in AI-generated news content has profound consequences. First, it makes fact-checking difficult, as readers and journalists cannot trace the sources or verify whether AI-generated information is based on credible references or misinformation. Second, it raises ethical concerns—without disclosure of how AI selects, filters, or prioritizes information, there is a risk of hidden biases influencing news narratives. For example, an AI model trained on biased datasets may reinforce stereotypes without clear accountability. Third, it reduces public trust in AI-driven journalism, as audiences tend to be skeptical of content whose origins and accuracy cannot be easily verified.

Most countries have yet to implement regulations specifically addressing AIGC transparency, making it a pressing issue for the future of journalism. If AI-generated news is intended to gain credibility, developers and policymakers must work together to create transparency standards, such as mandatory disclosure of data sources and algorithmic decision-making processes.

5. Conclusion

The credibility of news content is influenced by various factors. Compared with traditional news content, AI-generated news content has its own advantages and disadvantages in terms of credibility. From the perspective of advantages, massive training data increase the accuracy and relevance of AIGC, allowing more comprehensive coverage and a more nuanced understanding of events. In addition, the use of rigorous and objective AI algorithms prevents AI-generated news content from being influenced by the personal bias or selfishness of journalists, improving overall reliability and credibility. From the perspective of disadvantages, the inaccurate, false, or biased information contained in the database greatly increases the possibility of biased or untrue news content being produced, which is fatal to the credibility of AIGC. Moreover, the lack of information transparency in AI-generated news content, such as obscuring the algorithms used and training data due to technical problems, makes it unable to establish a trust bridge with readers such as what traditional news content does.

Currently, AIGC technology is developing at an unprecedented pace, playing an increasingly important role in journalism. Progress and problems are always inseparable. In the future, ongoing research on technological innovation is essential for AI-generated news content to solve its current challenges and achieve greater credibility. In addition, how to preserve the advantages of traditional news content and integrate it with AI-generated news content is also an important research topic for future journalism.

References

- Agovino, M., Carillo, M. R., & Spagnolo, N. (2022). Effect of media news on radicalization of attitudes to immigration. *Journal of Economics, Race, and Policy*, 5(4), 318-340. <https://doi.org/10.1007/s41996-021-00091-4>
- Al-Khalifa, H., AlOmar, T., & AlOlyyan, G. (2024). Natural language processing patents landscape analysis. *Data*, 9(4), Article 52. <https://doi.org/10.3390/data9040052>
- Andrada, G., Clowes, R. W., & Smart, P. R. (2023). Varieties of transparency: Exploring agency within AI systems. *AI & Society*, 38(4), 1321-1331. <https://doi.org/10.1007/s00146-021-01326-6>

- Fang, X., Che, S., Mao, M., Zhang, H., Zhao, M., & Zhao, X. (2024). Bias of AI-generated content: An examination of news produced by large language models. *Scientific Reports (Nature Publisher Group)*, 14(1), Article 5224. <https://doi.org/10.1038/s41598-024-55686-2>
- Han, J., Li, Q., Xu, Y., Zhu, Y., & Wu, B. (2024). Design of a trusted content authorization security framework for social media. *Applied Sciences*, 14(4), Article 1643. <https://doi.org/10.3390/app14041643>
- Johnson, J. A., Moore, B., Hwang, E. K., Hickner, A., & Yeo, H. (2023). The accuracy of race & ethnicity data in US based healthcare databases: A systematic review. *The American Journal of Surgery*, 226(4), 463-470. <https://doi.org/10.1016/j.amjsurg.2023.05.011>
- Miller, R. (2015, January 29). *AP's "robot journalists" are writing their own stories now*. The Verge. <https://www.theverge.com/2015/1/29/7939067/ap-journalism-automation-robots-financial-reporting>
- Nyilasy, G. (2020, April 22). *Fake news in the age of COVID-19*. University of Melbourne. <https://fbe.unimelb.edu.au/newsroom/fake-news-in-the-age-of-covid-19>
- Xie, Y., Zhang, Z., Liu, Y., Chen, E., & Li, N. (2024). Evaluation method of IP geolocation database based on city delay characteristics. *Electronics*, 13(1), Article 15. <https://doi.org/10.3390/electronics13010015>
- Zednik, C. (2021). Solving the black box problem: A normative framework for explainable artificial intelligence. *Philosophy & Technology*, 34(2), 265-288. <https://doi.org/10.1007/s13347-019-00382-7>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).