

Opinion

Digital transformation in journalism: Navigating big data governance and ethical challenges

Xinyang Li¹, Kunpeng Sun^{2,*}

¹ Tsinghua University, Beijing 100084, China

² Central University of Finance and Economics, Beijing 100081, China

* Corresponding author: Kunpeng Sun, sunkunpeng@cufe.edu.cn

CITATION

Li X, Sun K. Digital transformation in journalism: Navigating big data governance and ethical challenges. *Insight - News Media*. 2024; 7(1): 655.
<https://doi.org/10.18282/inm655>

ARTICLE INFO

Received: 3 September 2024
Accepted: 1 November 2024
Available online: 20 November 2024

COPYRIGHT



Copyright © 2024 by author(s).
Insight - News Media is published by PiscoMed Publishing Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license.
<https://creativecommons.org/licenses/by/4.0/>

Abstract: Digital transformation is reshaping journalism, driven by advancements in big data analytics, artificial intelligence (AI), and digital platforms. These technologies offer new opportunities for dynamic storytelling, audience engagement, and content personalization. However, they also present challenges related to data governance, privacy, and ethical standards. As news organizations increasingly rely on data-driven practices, they must navigate complex regulatory landscapes and address ethical dilemmas to maintain credibility and public trust. This paper proposes an integrated framework combining big data governance with journalistic ethics and examines empirical case studies of leading news organizations to illustrate the key challenges and strategies for ensuring responsible and effective use of data.

Keywords: digital transformation; journalism; big data; audience engagement; data privacy; AI; ethics

1. Introduction

Journalism today is at a critical juncture. The rise of digital technologies—particularly big data analytics, AI, and digital platforms—is transforming how news is produced, distributed, and consumed. These tools provide news organizations with unprecedented capabilities to understand audience behavior, deliver personalized content, and enhance engagement. Yet, as journalism becomes more data-driven, it also faces new challenges. Issues of data privacy, ethical standards, and compliance with regulatory frameworks are increasingly central to the sustainability and integrity of the profession.

This paper explores the impact of digital transformation on journalism, focusing on the governance of big data, ethical challenges, and the strategies that news organizations can employ to adapt to this new environment. We propose an integrated theoretical framework that combines big data governance principles with journalistic ethics, aiming to bridge the gap between technological innovation and ethical responsibility. By examining empirical case studies, we provide insights into how journalism can navigate the complexities of the digital age while maintaining its core values of transparency, accountability, and public trust.

2. The role of big data in modern journalism

2.1. Enhancing storytelling and audience engagement

Big data is fundamentally changing the nature of storytelling in journalism. By analyzing large datasets, news organizations can uncover patterns and trends that

traditional reporting methods might overlook. This data-driven approach allows for more in-depth investigative journalism, revealing hidden connections and insights that are critical to public interest stories. For instance, during the COVID-19 pandemic, news organizations used big data to track virus spread, provide real-time updates, and analyze the effectiveness of public health measures. Such data-driven reporting has been essential in keeping the public informed and engaged [1].

A notable example is The New York Times, which developed interactive maps and visualizations based on big data to illustrate the spread of COVID-19 across different regions. This not only enhanced storytelling but also provided readers with personalized information relevant to their locations [2].

Moreover, big data enables personalized news delivery, allowing media outlets to tailor content to the preferences and behaviors of individual users, thereby enhancing audience engagement. Personalized newsletters, AI-curated news feeds, and interactive data visualizations are just a few ways news outlets are using big data to connect with readers on a deeper level. The Guardian, for example, utilizes audience data analytics to understand reader preferences and to curate content that resonates with its audience, resulting in increased reader engagement and subscription rates [3]. These practices help not only in attracting larger audiences but also in retaining them by providing a more relevant and engaging news experience [4].

2.2. Data-driven journalism: Opportunities and risks

While big data offers many advantages, it also brings risks that must be carefully managed. The use of data analytics in journalism raises concerns about data quality, accuracy, and ethical use. The accuracy of data-driven stories depends on the quality of the underlying data, which may be biased or incomplete, especially when sourced from social media or third-party platforms. Misinformation can spread quickly if data is not properly vetted, and this can undermine public trust in journalism [5].

Furthermore, as news organizations collect more data on their audiences, they must also address privacy concerns. Audiences are increasingly aware of how their data is collected, stored, and used. Failing to be transparent about these practices can lead to distrust and reputational damage. The Cambridge Analytica scandal, involving data misuse from Facebook, heightened public awareness and concern over data privacy, affecting how audiences perceive data practices in journalism [6]. Implementing robust data governance frameworks that include regular audits and compliance checks can help mitigate these risks and ensure responsible data use [7].

2.3. Methodology

This study employs a qualitative case study methodology to explore how leading news organizations address the challenges of big data governance and ethical considerations in the context of digital transformation. We selected prominent cases, including The Guardian and The New York Times, based on their innovative use of big data and AI technologies, as well as their public commitments to ethical journalism. Data were collected from official reports, academic articles, and credible news sources to provide a comprehensive analysis of their strategies and practices.

3. Ethical challenges in the age of digital journalism

3.1. Navigating algorithmic bias and fairness

The increasing use of AI and algorithms in journalism presents ethical challenges, particularly around bias and fairness. Algorithms used for content curation and recommendation can inadvertently reinforce existing biases, limiting the diversity of information that audiences are exposed to. For example, if an algorithm prioritizes click-driven content, it may marginalize important stories that are less sensational but more informative. This bias can create “filter bubbles,” where users are exposed primarily to content that reinforces their existing beliefs, potentially polarizing public discourse [8].

A practical example is Facebook’s news feed algorithm, which has been criticized for promoting sensationalist content and creating echo chambers [9]. This has significant implications for journalism, as news organizations distributing content through such platforms may inadvertently contribute to the problem.

To address these issues, news organizations must adopt transparent and accountable AI practices. This includes regularly auditing algorithms for bias, diversifying data sources, and ensuring that AI-driven content decisions are aligned with editorial standards and ethical guidelines. The New York Times has begun developing “explainable AI” models, which provide transparency into how algorithms make decisions, helping to build trust with their audiences [10]. By doing so, they can help prevent the spread of misinformation and maintain the integrity of their journalism. Some organizations have begun developing “explainable AI” models, which provide transparency into how algorithms make decisions, helping to build trust with their audiences [11].

3.2. Protecting privacy in a data-driven world

As journalism becomes more reliant on data, protecting user privacy becomes paramount. News organizations must balance the need for data to drive engagement and personalization with the ethical obligation to protect their audience’s personal information. Implementing strict data protection measures, such as encryption and data minimization, is crucial to safeguarding privacy. Additionally, transparency about data practices, including clear consent protocols and data usage explanations, can help maintain trust and compliance with international laws like the GDPR [5].

Moreover, news organizations should consider adopting privacy-by-design principles, embedding data protection into the very fabric of their operations and technology. This proactive approach ensures that privacy is maintained at all stages of data handling, from collection to processing and storage, reinforcing trust in digital journalism [7].

4. Strategies for adapting to the digital age

4.1. Building robust data governance frameworks

To effectively manage the challenges of data-driven journalism, news organizations must develop robust data governance frameworks. These frameworks

should encompass data privacy, security, and ethical standards that align with both domestic and international regulations. Regular audits, transparent data use policies, and dedicated data protection officers can help ensure compliance and build public trust [11].

News organizations should also prioritize training for their staff on data ethics and security to build a culture of accountability and transparency. Providing journalists and editors with the tools and knowledge to handle data responsibly is essential for maintaining credibility in the digital age. The BBC, for instance, has implemented comprehensive training programs on data ethics for its employees [12]. Additionally, collaborating with other media outlets and industry bodies to develop and share best practices can further enhance data governance standards [13].

4.2. Leveraging technology responsibly

Leveraging technology responsibly is key to navigating the digital landscape. AI can be used to automate repetitive tasks, enhance content personalization, and improve operational efficiency. However, news organizations must ensure that AI tools are deployed transparently and ethically, with safeguards to prevent misuse. Blockchain technology, for instance, can be used to provide transparent records of digital transactions, ensuring that data is handled in a secure and accountable manner. For example, blockchain technology can be used to provide transparent records of digital transactions, ensuring that data is handled in a secure and accountable manner [14].

Moreover, adopting AI ethics frameworks that involve stakeholders in the decision-making process can help ensure the ethical use of technology in journalism [15]. The Ethical OS Toolkit, developed by the Institute for the Future, is an example of a resource that organizations can use to anticipate and address ethical risks associated with new technologies [16].

4.3. Engaging with audiences and building trust

In the digital age, building trust with audiences is more important than ever. News organizations must engage openly with their readers, providing transparency about their data practices and offering opportunities for feedback. Interactive platforms, such as comment sections, social media forums, and virtual town halls, can foster direct communication and dialogue between news outlets and their audiences. By being responsive to audience concerns and feedback, news organizations can strengthen their relationship with their readers, building loyalty and credibility [17].

The Guardian's "Reader-funded journalism" model is an example of engaging with audiences by involving them directly in funding and supporting journalism, thereby increasing transparency and trust [3].

Moreover, fostering a participatory culture where audiences contribute to content creation, through citizen journalism initiatives or crowdsourced investigations, can enhance engagement and diversify perspectives in news coverage. Such practices not only build trust but also enrich the content and relevance of journalism in a rapidly changing world [13].

5. Conclusion

Digital transformation not only offers significant opportunities for innovation in journalism but also presents substantial challenges in data governance, privacy and ethics. This paper has proposed an integrated framework that combines big data governance with journalistic ethics, contributing a novel perspective to the literature. Through empirical case studies of leading news organizations, we have illustrated the complexities and practicalities of implementing responsible data practices.

Our study highlights the need for robust data governance frameworks, the ethical use of technology, and transparent engagement with audiences to navigate the digital landscape effectively. The innovations of this study lie in the integration of theoretical and practical insights, offering actionable strategies for news organizations.

However, there are limitations to our research. The scope of case studies is limited, and further quantitative analysis could strengthen the findings. Future research could explore longitudinal studies on the impact of data governance frameworks on audience trust and investigate the role of AI ethics in journalism more deeply.

By addressing these challenges proactively, news organizations can ensure their continued relevance and integrity in the digital age, maintaining public trust and upholding the core values of journalism.

Acknowledgments: The authors would like to express their gratitude to the Fundamental Research Funds for the Central Universities for their support in the development of this Opinion.

Conflict of interest: The authors declare no conflict of interest.

References

1. Fernández-Rovira C, Álvarez Valdés J, Molleví G, et al. The digital transformation of business. Towards the datafication of the relationship with customers. *Technological Forecasting and Social Change*. 2021; 162: 120339. doi: 10.1016/j.techfore.2020.120339
2. The New York Times. Coronavirus map: Tracking the global outbreak. The New York Times; 2020.
3. Guardian Media Group. Annual report and financial statements 2021. Guardian Media Group; 2021.
4. Carrion KYP, Coronel-Salas G. The digital transformation of Ibero-American journalism. 2020 15th Iberian Conference on Information Systems and Technologies (CISTI). 2020; 1-6. doi: 10.23919/cisti49556.2020.9141118
5. Schneider S, Kokshagina O. Digital transformation: What we have learned (thus far) and what is next. *Creativity and Innovation Management*. 2021; 30(2): 384-411. doi: 10.1111/caim.12414
6. Isaak J, Hanna MJ. User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*. 2018; 51(8): 56-59. doi: 10.1109/mc.2018.3191268
7. Vial G. Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*. 2019; 28(2): 118-144. doi: 10.1016/j.jsis.2019.01.003
8. Gong Y. Thoughts on the application of big data in various fields. *Data Science Journal*. 2018.
9. Pariser E. *The filter bubble: What the internet is hiding from you*. Penguin Press; 2011.
10. Hansen M, Roca-Sales M, Keegan JM, & King G. *Artificial intelligence: Practice and implications for journalism*. Tow Center for Digital Journalism, Columbia Journalism School; 2017.
11. Molchanova S. Impact of digital transformation on the strategic renewal of companies. *Ekonomika i Upravljenje: Problemy, Resheniya*; 2021.

12. BBC Academy. Data journalism and data ethics training modules. BBC Academy; 2020.
13. Kosterich A, Weber MS. Transformation of a Modern Newsroom Workforce. *Journalism Practice*. 2018; 13(4): 431-457. doi: 10.1080/17512786.2018.1497454
14. Ølnes S, Ubacht J, Janssen M. Blockchain in government: Benefits and implications of distributed ledger technology for information sharing. *Government Information Quarterly*. 2017; 34(3): 355-364. doi: 10.1016/j.giq.2017.09.007
15. Warner KSR, Wäger M. Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*. 2019; 52(3): 326-349. doi: 10.1016/j.lrp.2018.12.001
16. Institute for the Future. The ethical OS toolkit: Anticipate the future impact of today's technology. Institute for the Future; 2018.
17. Robinson S, Lewis SC, Carlson M. Locating the "Digital" in Digital Journalism Studies: Transformations in Research. *Digital Journalism*. 2019; 7(3): 368-377. doi: 10.1080/21670811.2018.1557537