Development Opportunities and Breakthrough Paths of Data Journalism in the Context of AIGC

Qianhui Jia

Chengdu University, Chengdu, 610000, China

Jiaqianhui3350507@163.com

Abstract. Generative AI provides multiple innovative paths for the development of data journalism, empowering all aspects of data journalism, from automated news production, enhanced data visualization to personalized news generation, cross-media integration. Generative AI has great potential for application in the data journalism industry, but it also brings new challenges such as data ethics and news authenticity. Journalists and institutions need to take the initiative to break through the bottleneck of data journalism development through innovation in automated data journalism production, enhanced data visualization and interactive experience, data-driven in-depth analysis, and cross-media integration. In order to promote the high-quality development of data journalism.

Keywords: Data Journalism; AIGC; Innovation-driven development.

1 Introduction

Data Journalism is also called Data Driven Journalism. As the world's first book on data journalism business, data journalism Manual points out that "data journalism creates new possibilities for combining traditional news sensitivity and persuasive narrative ability with massive digital information". Data news has changed the traditional news production model and the entire operational system."[1]. In a narrow sense, data journalism is "a new way of news reporting based on data capture, mining, statistics, analysis and visual presentation"[2]. With the rise of artificial intelligence, the traditional production mode of data journalism began to be difficult to meet the development needs. The effective use of artificial intelligence generated content by data journalism not only improved the efficiency and quality of data journalism production, but also innovated the narrative mode of data journalism, and brought more development opportunities and traffic for news organizations[3]. China's data journalism started relatively late, and in the process of development gradually exposed some phenomena, such as the lack of innovation in content, the release of time to be strengthened, and many other problems to be solved[4]. In the face of the development dilemma of data journalism, it is necessary to further explore the impact and role of artificial intelligence technology in the development of data journalism, to promote the

F. Zeng et al. (eds.), *Proceedings of the 2024 7th International Conference on Humanities Education and Social Sciences (ICHESS 2024)*, Advances in Social Science, Education and Humanities Research 887,
https://doi.org/10.2991/978-2-38476-323-8_65

effective dissemination of data journalism, and to promote the innovative development of data journalism.

2 The Main Features of the Development of Data Journalism

2.1 Low Utilization of the Mobile Platform and a Single form of Visualization

According to CNNIC's 54th Statistical Report on the Development of China's Internet, as of June 2024, the number of Internet users in China is nearly 1.1 billion, and the Internet penetration rate is 78.0%. Mobile terminals such as mobile phones have become the main channels for Internet users to obtain news information. However, data journalism on mobile terminals are mostly in the form of traditional news reports, and data journalism reports are updated less frequently, The number of releases is small, mainly static information charts, and the use of dynamic interactive forms is insufficient. In terms of specific visual style, Xinhua News Agency presents diversified forms of expression according to the platform attributes and news types, Data maps and other rich forms are used.

2.2 The Concept and Process of News Production have Not yet Realized the Transformation of Data

The current production of data journalism is mainly based on traditional news reports, the data source is mainly the data released by the government, the website or policy documents are the main source of data journalism to obtain data, the data source is relatively single, the limitations of data channels, affecting the breadth and depth of data journalism reports. The most important impact of the emergence of data journalism is the impact on the traditional concept of news production, data reporter Mirko Lorenz has put forward the four steps of data journalism production, that is, mining data, filtering data, data visualization, and the completion of the production of news reports[5]. However, in the current data journalism reports, there is still a lack of attention to data, mining, analysis and interpretation, and the content is mainly a "secondary processing" of traditional news reports, mostly a copy of the data material, with limited depth of data mining.

2.3 There is an Urgent Need to Upgrade Data Thinking

The core of data journalism lies in the use of data, and the current data journalism focuses on the visual beautification and visualization of data, rather than the analysis and interpretation of data, and data, as the main news production factor, has not yet been strengthened as the main body of the narrative. The key to the influence of data journalism lies in re-examining the role of data in news production, data as the main

source of information and narrative source, expanding data sources, mining news value in data, and the use of data thinking are the key and difficult points in data journalism.

2.4 Rational Adherence to Technology-Driven forms of Visualization

With the new round of technological revolution represented by VR, AI, etc., data journalism reports also have many innovations AI painting, audio posters, micro videos, small games and other news products continue to brush the screen. In the face of the impact of new technologies, data journalism reports still take graphic news as the main form. Behind the bottleneck period of data journalism development, how to seize a new round of transformation opportunities and use emerging technologies to carefully plan data journalism media products has become an important problem to be solved. The goal of data visualization depends on the data. Different types of data require different visual designs, and also determine visual choices such as visual styles and symbols.

2.5 The Mobile Bias in the Trend Towards Platform Consolidation

Under the new media environment, the media's communication channels tend to be diversified. At present, major media mainly publish data journalism reports at both the PC end and the mobile end, focusing on the traditional PC end to carry out special reports. The layout of the mobile end is mostly the "secondary replication" of the PC end content, which has not yet formed the "ring breaking" effect of the platform. Faced with the reality that mobile phones are the largest Internet terminals, the mobile terminal has not yet become the main communication platform for data journalism reporting, the portability and high viscosity advantages of the mobile terminal have not yet been fully utilized, and the effective access of data journalism in a wider range has not yet been achieved.

3 The Impact of Generative Artificial Intelligence on Data Journalism

3.1 Automated Data Processing and Analysis

Traditional data journalism production requires manual collection and screening of data. The emergence of generative AI can automatically extract information from public data sets, policy documents, social media, and databases, and instantly sort out data trends and patterns, such as:, AI can quickly scan thousands of pages of government documents or financial reports, extract key information, and carry out a large number of data collection and preliminary analysis, which greatly saves the time for data collection and sorting. At the same time, AI can cover data sources that were difficult to reach before, and automatically complete data cleaning, duplication removal, classification and other tedious work, providing cleaner More structured data information. On the other hand, generative AI can predict and analyze data, model data,

predict future trends through machine learning algorithms, and provide more in-depth insights. For example, it can help analyze climate data to predict the long-term impact of climate change on a place. Generative AI's empowerment of data journalism enhances the depth and breadth of data mining.

3.2 Enhancing the Efficiency of News Production

Generative AI is innovating the way of news writing, especially in the field of data-driven news. It can automatically generate data journalism types based on fixed formats, especially in fields rich in quantitative data such as financial reports, sports events and weather forecasts, AI's application in this scenario is very efficient. It can automatically generate accurate and logical news releases from data to improve the efficiency of data journalism reporting. At the same time, AI can generate multiple versions of news reports based on the same set of data, adjust the language style, depth and content to adapt to different audiences, and promote more personalized and diversified data journalism. The application of generative AI deepens the data-driven narrative. According to generative news, with the assistance of AI, the most newsworthy data points can be identified, and these data can be organized in the form of stories. Data driven narrative reports can be written according to complex data sets. For data journalism, this means that journalists can focus more on interpreting data rather than mechanically writing press releases.

3.3 Innovations in Data Visualization

Generative AI can help create more intuitive, dynamic and interactive visual works, make data journalism more vivid and easy to understand, and improve the timeliness of data charts. Traditional data journalism charts are mostly made of static line charts, bar charts, maps, etc. Generative AI can intelligently select the most appropriate visual form according to the type and characteristics of data, helping to build dynamic The real-time updated chart automatically adjusts according to the latest data and reflects the latest results in real time to realize real-time dynamic visualization. On the other hand, generative AI can provide a higher level of interactive experience and data narrative, allowing readers to explore data on their own, It helps readers understand the key information of charts. This interactive form not only increases the sense of user participation, but also makes news reports more attractive, data journalism more personalized, and news intuition and user experience.

3.4 Customized News Generation that

Generative AI not only changes the way of data journalism production, but also automatically generates personalized and customized data journalism content according to the interests and reading habits of readers. According to the needs and preferences of different users, different versions of data journalism reports are generated to provide personalized news experience. For example, localized data journalism reports are generated based on specific data in the user's region, Analyze

users' browsing behavior, reading history and interest preferences, customize news content for them, and make personalized recommendations. In addition, generative AI can generate multilingual data journalism reports, automatically translate and localize content, thus expanding the global audience of data journalism, so that users from different languages and cultural backgrounds can enjoy the same data journalism service.

4 Risks of Generative Artificial Intelligence in the Production of Data Journalism

Generative AI brings opportunities for the development of data journalism, but there are also certain risks and challenges, mainly related to news authenticity, ethical issues, transparency, and the long-term impact on the news industry.

4.1 Misuse and Misinterpretation of Data

Generative AI relies on data for news generation, but cannot judge the quality and source of the input data. Therefore, when the input data has problems such as deviation, incompleteness or false data, the data journalism content generated by AI will have misleading results. At the same time, generative AI can process a large amount of data, but it cannot always accurately understand the background and complexity, This will lead to the generated data journalism works ignoring the context behind the data, incorrectly deriving causality, and leading to wrong narrative and misleading reports.

4.2 Ethical Responsibility and Algorithmic Bias

There are also certain risks in ethical disputes and responsibility attribution. If there are data errors or misleading in AI generated data journalism works, it will lead to legal and ethical disputes. At the same time, the use of generative AI will reduce the transparency of data journalism production, weaken the credibility of news to a certain extent, and affect readers' trust in news organizations.

Since AI model mainly relies on training data, these data may form algorithm bias. If AI generates news based on biased data, it may aggravate social injustice or unbalanced communication narrative to some extent.

4.3 Disinformation and News Homogenization

Although generative AI performs well in the speed and efficiency of generating content, it is also easy to promote the spread of false information or misleading content, including the automatic generation of false information charts, resulting in the proliferation of false data content. In addition to text news, generative AI can also generate multimedia content such as fake video and audio, further blurring the boundaries between authenticity and forgery, This capability may be maliciously used to create and disseminate false data journalism, which will have a serious impact on

public opinion. In addition, the way AI generates data journalism is mainly based on data and algorithms, which easily leads to the homogenization trend of data journalism content. A large number of automatically generated data journalism content tends to lack depth, ignoring the narrative perspective, cultural background and human perception, and weakening the diversity of data journalism reports.

4.4 Lack of in-Depth Analysis of Complex Issues

Generative AI is good at processing a large amount of structured data, and can generate concise news content based on existing data. However, there are certain shortcomings for complex, emotional or in-depth analysis issues. For complex topics, most of the data journalism generated by AI involve superficial problems, and it is difficult to capture the unique insights of human journalists through in-depth interviews and research The production of data journalism may lead to some important social issues not being fully discussed or oversimplified. Based on the unique characteristics of data journalism, the narrative theme status of data, and the dependence on data, it is easy to lead to a lack of human perception, emotion and social responsibility.

5 Innovative Paths for the Development of Data Journalism

In the context of generative AI, data journalism has more innovation paths and development potential. generative AI not only improves the production efficiency of data journalism, but also brings new possibilities for news content, news form and news experience innovation.

5.1 Innovations in Automated Data Journalism Production

Generative AI can enable the automation of data journalism production. With generative AI, news organizations can generate real-time data journalism more quickly, especially in fields based on quantitative data, such as economic data, meteorological monitoring, etc., by accessing real-time data sources, AI can automatically generate and update news content to ensure that audiences always get the latest news information. Because AI can generate a large number of data journalism on different topics at a lower cost, news organizations can more effectively meet the data journalism needs of various segments. In addition, AI can be preset by setting rules or conditions, Automatically generating specific types of data journalism reports based on basic data when a specific event occurs can not only improve the response speed of news reports, but also enhance the timeliness and accuracy of data journalism.

5.2 Enhancing Data Visualization and Interactive Experience

The generative AI promotes the innovation of data visualization and provides a more attractive and interactive presentation mode for news. On the one hand, generative AI

can automatically generate dynamic charts and interactive visualization tools based on data, such as trend charts, maps, relationship network diagrams, etc. These tools can help audiences understand complex data more intuitively. On the other hand, they can promote interactive news to become an important innovation field of data journalism. Traditional data journalism is mostly static charts, Through generative AI, high-quality interactive charts, maps and data tools can be produced, and customized visual news can be generated according to the interests and needs of each user. The audience can select specific parameters through the interactive interface, generate data graphs that meet their own interests, and obtain personalized insights through their own exploration, It can further enhance the personalized experience and user stickiness of data journalism.

5.3 Data-Driven in-Depth Analysis that

Through the empowerment of generative AI, data journalism can further dig out the hidden laws and trends in complex data, and provide a more in-depth analysis perspective for news reporting. First, predictive data journalism. Through machine learning models, AI can predict possible future news events based on past data trends, and provide forward-looking content for the audience; Second, data investigation and revealing reporting. AI can automatically mine abnormal patterns from big data to help reveal hidden problems. Through AI's large-scale data processing capability, it is helpful to quickly find clues of news value, so as to conduct more in-depth investigation; The third is to conduct correlation data analysis. Generative AI can conduct correlation analysis on data in different fields, find potential cross domain connections, and provide more valuable reporting perspectives for in-depth reports.

5.4 Innovations in Cross-Media Convergence

Generative AI can not only generate text, but also automatically generate audio and video. Generative AI can help data journalism integrate multiple media forms and provide more rich cross media news experience. The innovative form of multimedia integration can meet the preferences of different audiences for news content forms. In addition, AI combines virtual reality, augmented reality and other technologies, Data can be displayed to the audience through three-dimensional models and other forms, leading the audience to "walk into" the news event scene, and more intuitively feel the data and facts in the news. This immersive experience can not only enrich the expression form of data journalism, but also increase the attractiveness of data journalism content.

6 Conclusions

At present, there are problems such as low number of data journalism reports, lack of visualization form, single data mining, etc., which are in the bottleneck period to be broken through. Generative AI provides a variety of innovative paths for data

journalism, empowering all aspects of data journalism from data collection to news writing to visualization and personalized news production. Generative AI is greatly changing the production and consumption of data journalism, promoting the industry to re-examine the importance of data as the main element of news production, strengthening "data thinking" and innovating data visualization. Use generative AI to empower data journalism production and make full use of the important advantages of mobile platform to promote the realization of data journalism bottleneck period To create more attractive and influential data journalism works.

References

1. Chen Hong, Qin Jing The History, Current Status, and Development Trends of Data News [J] Friend of the Editor, 2016, (01): 69-75.
2. Zhanangchun ,31(05):25-33.Fang Jie, Yan Dong. "data journalism" from a Global Perspective: Concept and Practice [J]. International Press, 2013,06:73-83.
3. Wang Xiaojun. Synergy and Symbiosis: Application Analysis of AIGC in data journalism [J]. Science and Technology Communication, 2024, 16 (11): 1-5. DOI: 10.16607/j.cnki.1674-6708.2024.11.005.
4. Ma Li. Research on the Development Dilemma and Path of data journalism in the New Media Era [J]. Western Radio and Television, 2024,45 (08): 21-24.
5. Liu Yikun data journalism production in the era of big data: status quo, impact and reflection [J] Modern Communication (Journal of Communication University of China), 2014,36 (11): 103-106.