

The Role Of Data Journalism In Enhancing Public Understanding Of Complex Issues

Peng Chengying¹

¹Kangwon National University Master Student, Graduate school of convergence, Kangwon National University Chuncheon Campus, 1 Kangwondaehakgil, Chuncheon-Si, Gangwon State, 24341 Republic Of Korea. 13719592813@163.com

Abstract

This study investigates the role of data journalism in fostering public understanding of complex social issues, through an analysis of audience preferences, trust and journalists' organisational challenges. Using a descriptive research design based on purposive sampling of 400 respondents (300 news consumers and 100 journalists), the findings indicate that data journalism is valued most in reporting on public health (34%) and environmental issues (26%). This emphasis highlights the significance of data journalism where factual reporting and transparency are a prerequisite. 64% of consumers reported high trust in that journalism promotes increased trust in news media by presenting empirical evidence and statistics, reinforcing the potential of data journalism to rebuild media credibility. Not surprisingly, consumers preferred infographics (39%) and interactive dashboards (28%) pointing to an overall demand for visually driven and accessible reporting. On the newsroom side, interactions revealed that barriers to data driven reporting were as follows: lack of technical skills (37%) and limited access to quality and/or trusted data (28%) to generate appropriate data driven socio-economic reports. There is support for data journalism in terms of civic engagement and media trust as long as barriers to data journalism include institutional training, resource allocation and policy/legislative infrastructure are removed.

Keywords: Data Journalism, Public Trust, Visual Storytelling, Media Literacy, Journalistic Barriers, Public Engagement.

1. INTRODUCTION

In an information-rich landscape that tends to be fragmented, the ability of the public to comprehend and engage with complex public concerns (like climate change, economic inequality, public health crises, and elections) is increasingly contingent upon the quality and clarity of journalism. The older storytelling forms are valuable for understanding important events or issues, but, in general, they often do not provide the full scope of complex problems that can involve large amounts of data, many layers of patterns, or very technical information. This phenomenon laid the foundation for the development of data journalism - a new form of journalism that provides context around raw data and includes presentation methods like data visualization, statistical methods, and investigative reporting helping to transform raw information into meaningful narratives. Data journalism, as a form of journalism provides the initial factual accuracy of journalism, and adds interaction and visual tools to communicate these data, using interactive maps and visual tools to break down very technical content for broader audiences. Data journalism provides context to abstract data for human understanding; it links numbers to lived experience. The growing importance of data journalism, especially during a time of misinformation, low trust in media, and increased public demand for transparency, has only increased. Journalists can use data as a storytelling device to shed light on systemic issues, elucidate hidden patterns that underpin them, and provide evidence-based information to improve civic literacy. Furthermore, data journalism has a critical role in accountability journalism which empowers citizens to question, check, and act on policy decisions and public issues with an informed knowledge base. This project is intended to illuminate how data journalism can help deepen the public's understanding of the complexities of issues through audience perceptions, audience consumption habits, and the challenges facing journalists with data driven reports. This understanding is important for the future of journalism, and for bolstering democratic engagement through informed public discourse.

2. LITERATURE REVIEW

Carlson (2018) examined the rise of automated journalism—also described as robotic reporting—through a study of labor practices in nine news organizations and their content sourcing practices. He focused on how automation has changed both labor practices and layout formats of news content and the authority of journalists. He noted that automation may speed up content production, however, concerns arise from the lack of human editorial judgment, value in traditional journalistic roles, and resulting ineffectiveness (e.g., losing the value of a journalist's role). Carlson's research provided the initial overview of how data-driven and algorithmically produced news content is rewriting the boundaries of journalistic identity and labor relations as established in the big data era.

Dash et al. (2019) focused on using big data analytics in the healthcare sector exploring the role of data management, predictive modeling, and evidence-based decision-making. While not strictly a journalism publication, it demonstrated how large datasets could be applied to generate meaningful insights and advance public health outcomes. The authors addressed future possibilities for managing structured and unstructured health data and highlighted the continued need for coherent data governance frameworks. The research also highlighted the role of interdisciplinary collaboration to manage complex data which is relevant to data journalists who work in health communication and reporting in times of crisis like the COVID-19 pandemic.

Georgiadou and Matsiola (2023) studied how journalism students conceptualized data journalism as a skill and type of professional discipline, as well as the extent to which educational interventions can improve awareness and interest. Their research found that students appreciated the increasing relevance of data journalism in the media industry, but also noted that they felt unprepared since they had limited exposure to the curriculum and training, and limited technical training. They underscored the importance of journalism education addressing data literacy, data analytics, coding and visualization tools to better meet the expectations of the pattern of the profession. Their study provided knowledge essential for understanding future writers and data journalists and it also demonstrated the imperative need for pedagogical reform to address the need for data driven reporting.

Gil-Garcia et al. (2018) explored how digital governance and public management are vis-a-vis, with increasing reliance on data and information technologies as fundamental components of new governance models. They stressed the significance of data integration and transparency on basic levels of public service provision and accountability. While their focus was on public administration, the results of their report are relevant for data journalism, especially their intention for the use of open data and the capabilities of journalists who can make sense of and narrate complex public data. The report suggested building networks of collaboration between government groups and data professionals, including journalists, to advance the purpose of making public data available and enabling usability for civic action and news media coverage.

Hanitzsch and Vos (2018) challenged the conventional understandings of journalistic roles, suggesting that journalism is not only a function of the democratic watchdog role, but also grew into contexts related to societal and cultural roles too; and that there should be a more nuanced way of thinking about how journalists operate within the context of political and social life; their study illustrated the versatility of journalistic roles across cultures and various media systems and insisted we simultaneously keep data journalism in consideration within these shifting boundaries; participants felt as journalism moves to contain elements of data and storytelling, roles would become more significant defined as explainer, analyst, interpreter and so forth, and in relation to how the public comprehends this information in a more saturated and complex media ecosystem.

3. RESEARCH METHODOLOGY

This study adopted descriptive research design using purposive sampling of 400 respondents (300 consumers, 100 journalists). To gather data, structured online questionnaires were utilized. Data was analyzed using percentage distribution and visually presented via tables and graphs, produced on Excel and Google Forms.

3.1. Research Design

This study utilized a descriptive research design that aimed to understand public perceptions, preferences, and perceived impact of data journalism and barriers journalists experienced when practicing data journalism. This research design allowed for quantitative data to be collected to describe consumer and journalist perceptions, preferences, trends and relationships towards data journalism. Descriptive statistics such as percentages were used to analyze the distribution of responses and barriers in the main categories, i.e, awareness of data journalism, trust in data journalism, relevance of topics, delivery format preferences, perceived challenges.

3.2. Data Collection Method

Data were collected using a structured online questionnaire which included closed-ended questions and multiple-choice questions. The survey was shared widely on social media, email groups, and professional journalist networks to reach as many consumers and journalists as possible. The questionnaire contained two main sections; one for general news consumers, while the other targeted practicing journalists. Each section aimed to capture relevant data on awareness, perceptions, behaviors, and professional barriers.

3.3. Sampling Technique and Sample Size

A purposive non-probability sampling technique was used to purposively select individuals or groups who indicated that they were regular consumers of news media or who actively do journalism. In total, 400 respondents completed the survey, 300 of whom were general consumers and 100 who were professional journalists. Having this many respondents ensured we had balanced insights from both sides of the data journalism ecosystem—that of generally accepting public reception and of newsroom practices.

3.4. Data Analysis Tools and Techniques

The collected data were organized into tables and analyzed according to descriptive statistics, specifically percentage distribution. The descriptive statistics were reported using bar graphs/charts so readers could quickly and visually interpret and compare the data. Data interpretation addressed dominant trends and recognizable conclusions that aligned with the study's objectives. Data entry, analysis, and visualization were performed in Microsoft Excel and Google Forms.

4. Data Analysis And Interpretation

Table 1 shows the topic areas where people perceive the most value in data journalism, demonstrating that Public Health takes the top spot at 34% of consumer preference—with people perceiving data journalism to be most helpful in health topics (i.e., COVID-19), followed by Climate Change & Environment at 26%, indicating there is also strong public interest in data journalism in relation to environmental stories about data. Economy & Budgeting received 21%, indicating there is moderate interest in reporting the ways in which financial data is presented and explained. Crime & Justice was 11%, while Sports Analytics ranked the lowest at 8%, indicating that consumers view these areas as less important for data journalism.

Table 1: Topics Where Data Journalism Is Most Valued by the Public

Topic Area	Percentage of Consumer Preference (%)
Public Health (e.g., COVID-19)	34%
Climate Change & Environment	26%
Economy & Budgeting	21%
Crime & Justice	11%
Sports Analytics	8%

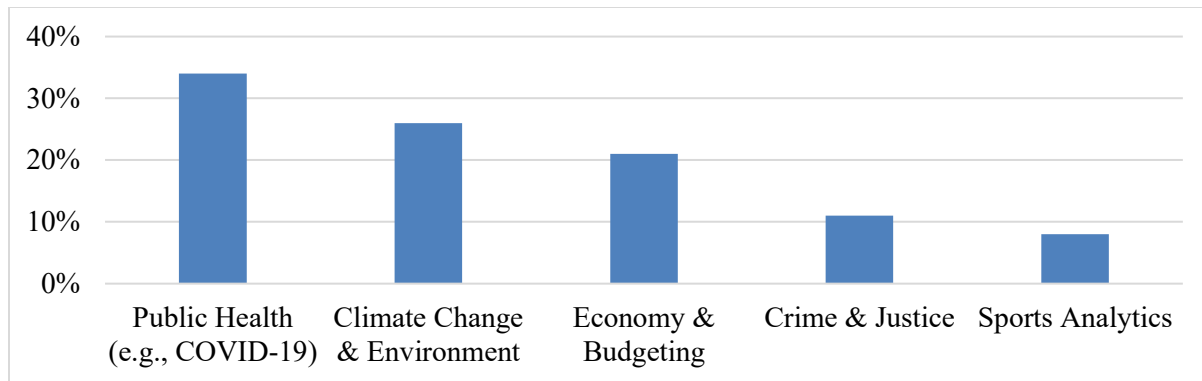


Figure 1: Graphical Representation of Topics Where Data Journalism Is Most Valued by the Public

The table indicates emphasis on public interest within socially impactful and knowledge-dominated rarefied spaces. The strong preference for data journalism in Public Health implies the respected utility in understanding possible context driven by transparent, factual, and data-fuelled communications and changing public behaviours in response to global public health crises such as COVID-19. Likewise, the interest in Climate Change & Environment suggests ever-increasing demand for clear, data-driven, and visually represented context to understand complexity in environmental anecdotes and challenges. Moderate interest in Economy & Budgeting refers to the need for clarity, especially in finance and related matters amidst wider economic concerns. The lower levels of interest in Crime & Justice and Sports Analytics suggest that while they are linked to data journalism, they are not seen as high importance or impact in comparison to health or environmental reporting. The general take away is that while appreciation of data journalism likely comes from public interest, they are reflected most when associated with relevant issues are are urgent to society.

Table 2 shows how data journalism has influenced public trust in news media. The data summarised above indicates that 41% of consumers say they have a Moderate Increase in Trust with the use of data journalism, the largest proportion. This is followed by 22% who claim to have their Trust Increased Significantly, which demonstrates that many people consider data journalism to have a positive impact on the credibility of news media. In comparison, 25% of respondents are with No Change in Trust, and 12% say they have a Decreased Trust. It may be worth noting that the 25% and 12% groups are relatively neutral and cynical respectively.

Table 2: Impact of Data Journalism on Public Trust in News

Level of Increased Trust	Percentage of Consumers (%)
Trust Increased Significantly	22%
Trust Increased Moderately	41%
No Change in Trust	25%
Trust Decreased	12%

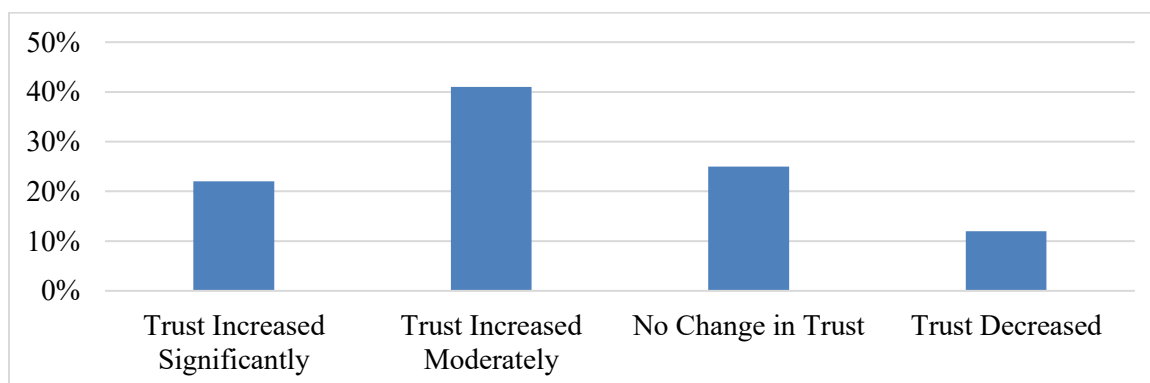


Figure 2: Graphical Representation of Impact of Data Journalism on Public Trust in News

The results show data journalism has an overall positive impact on increasing public trust in news with a combination of 63% of the respondent's reporting moderate or significant increase in trust. This suggests that the public is increasingly valuing evidence-based, transparent reporting which is the essence of data journalism. With 25% reporting no change suggests data journalism provides benefits to public trust, although there is not always an accompanying shift in perception. This may be due to pre-established biases or less skepticism towards certain media institutions. The 12% that reported a decrease in trust may have misinterpreted the data they engaged with, felt there was some manipulation of the data being presented, and/or simply, were unable to interpret or accurately engage with the data which suggested they lacked data literacy. Overall, the table demonstrates that data journalism, while more of less effective, is an important part of building news trust, and that public trust and its increase or decrease relates to personal consumption and expectations of news.

Table 3 outlines the main barriers journalists face in tracking data journalism. The most cited barrier is Lack of Technical Skills, with 37% of journalists reporting a definite skills gap in working with data analysis, coding or visualization. Access to Quality Data at 28% indicates challenges with finding accurate and up-to-date datasets that are publically available. Time Constraints in Newsrooms is cited by 19%, which recognizes the fast-paced nature of journalism and lack of time to dedicate to more thorough data work. Lack of Editor Support cited by 10%, shows there is organizational or managerial hesitation in developing a data journalism culture. Finally, High Cost of Tools/Software is reported by 6%, which signifies that despite being far from the most frequent barrier, cost is still a barrier, albeit, a lesser barrier.

Table 3: Major Barriers Faced by Journalists in Practicing Data Journalism

Barrier	Percentage of Journalists (%)
Lack of Technical Skills	37%
Limited Access to Quality Data	28%
Time Constraints in Newsrooms	19%
Lack of Editorial Support	10%
High Cost of Tools and Software	6%

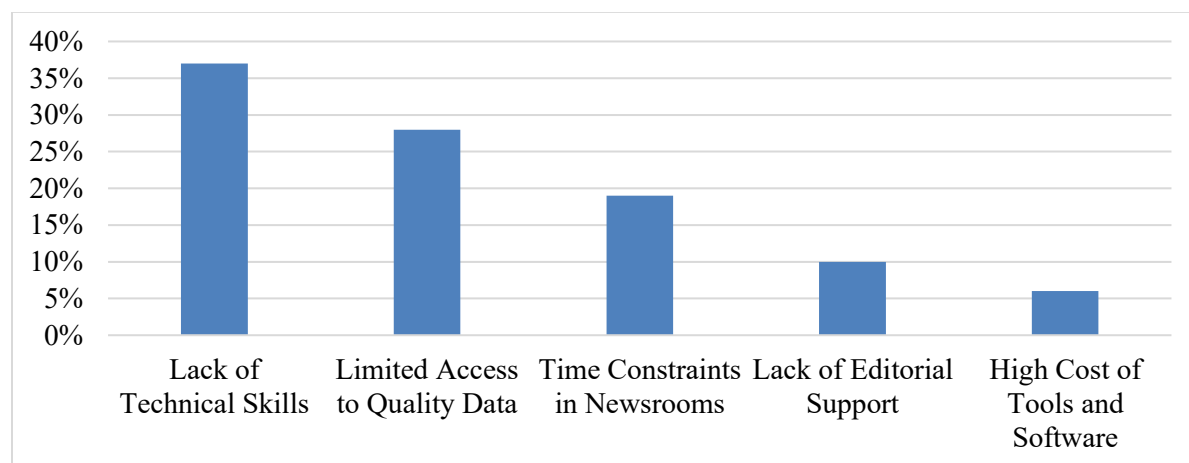


Figure 3: Graphical Representation of Major Barriers Faced by Journalists in Practicing Data Journalism

The data shows that the successful practice of data journalism is largely limited by structural and skill based barriers. The clear prevalence of Lack of Technical Skills highlights the urgent need for informed training and upskilling programs within newsrooms. The high percentage of respondents naming Limited Access to Quality Data reinforces the importance of transparency in data and open data initiatives of both public and private bodies. Time Constraints indicate that establishing day-to-day workflows that incorporate the steps needed for data journalism still present a logistical barrier to practice meaning that restructuring workflows for data journalism or dedicated teams may be warranted. The scores for Lack of Editorial Support and High

Cost of Tools reflect a need for institutional sustainability in both commitment and resources. Overall, removing these barriers is crucial for the widespread and effective use of data journalism in contemporary newsrooms.

In Table 4, consumer preferences for data journalism formats are displayed. The most preferred format is Infographics, with 39% of consumers expressing preference, suggesting a strong preference for visually stunning and straightforward information. Interactive Dashboards are the second format, with 28% of consumers showing a preference to data journalism formats that allow them to explore the data at their own discretion. Short Data-Driven Videos were the preferred format for 22% of respondents, indicating a sizeable preference for audiovisual representations of storytelling that include data elements. Lastly, only 11% of consumers preferred Long Form Reports with Charts, indicating a much clearer preference to less dense formats that include less text heavy analysis content.

Table 4: Preferred Formats of Data Journalism Among Consumers

Format	Percentage of Preference (%)
Infographics	39%
Interactive Dashboards	28%
Short Data-Driven Videos	22%
Long-Form Reports with Charts	11%

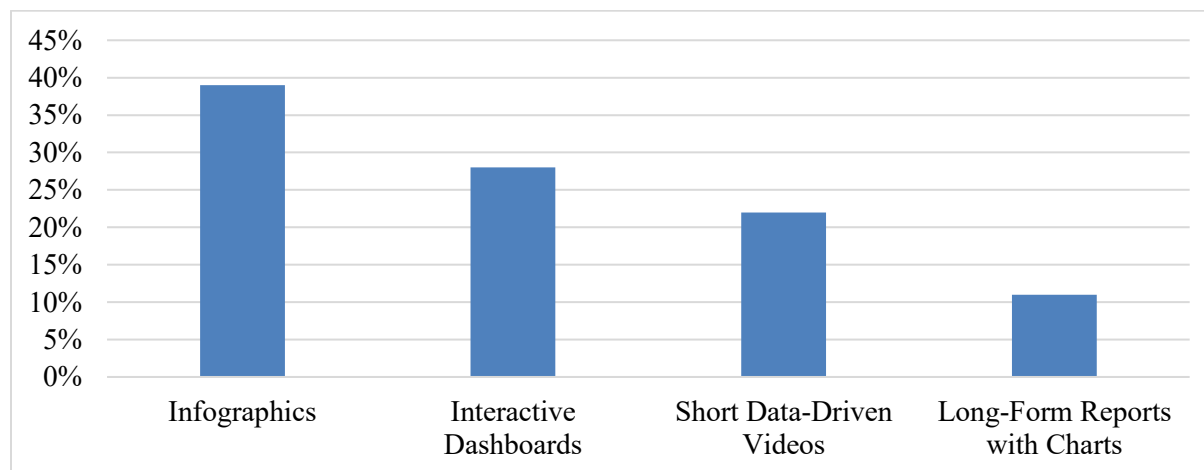


Figure 4: Graphical Representation of Preferred Formats of Data Journalism Among Consumers

The data has clearly indicated a consumer bias toward formats that are visually appealing and easy to understand. The preference for Infographics and Interactive Dashboards suggests that simplicity, clarity, and interactivity are the themes to consider in doing data journalism for mass audiences. These two formats allow for quick consumption of information and are likely attractive to not only casual readers but engaged readers as well. The moderate preference for Short Data-Driven Videos indicates that there is an interest in video content and video storytelling, especially in the digital space. At the same time, the relatively low preference for Long-Form Reports with Charts indicates that traditional data storytelling formats that demand more time and attention faced a challenge. These results suggest that in order to maximize reach and engagement, data journalism should prioritize visual and interactive methods of delivery that align with the media consumption habits of contemporary audiences.

5. CONCLUSION

This study's results highlight the significant role that data journalism takes to expand public understanding of complex issues like public health, climate change, and economic policy. Most consumers reported being more trustworthy and engaged when news is delivered through a clear, data-informed medium like infographics and interactive dashboards, showing a clear preference towards engagement at a distance

through accessible formats, particularly visual content. However, data journalism is not inherently simple to execute; journalists face barriers including limited technical skills, limited access to quality data, and lack of time in busy newsroom situations. By addressing barriers through institutional support, training, and open data, we can begin to harness data journalism as a tool to promote more informed public discourse, transparency, and accountability in the digital age.

REFERENCES

1. Boon-Itt, S., & Skunkan, Y. (2020). Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. *JMIR public health and surveillance*, 6(4), e21978.
2. Bulger, M., & Davison, P. (2018). The promises, challenges, and futures of media literacy. *Journal of Media Literacy Education*, 10(1), 1-21.
3. Carlson, M. (2018). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. In *Journalism in an era of big data* (pp. 108-123). Routledge.
4. Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of big data*, 6(1), 1-25.
5. Georgiadou, E., & Masiola, M. (2023). Understanding and enhancing journalism students' perception of data journalism. *Journalism and Media*, 4(4), 1232-1247.
6. Gil-Garcia, J. R., Dawes, S. S., & Pardo, T. A. (2018). Digital government and public management research: finding the crossroads. *Public management review*, 20(5), 633-646.
7. Hanitzsch, T., & Vos, T. P. (2018). Journalism beyond democracy: A new look into journalistic roles in political and everyday life. *Journalism*, 19(2), 146-164.
8. Head, B. W. (2022). *Wicked problems in public policy: Understanding and responding to complex challenges* (p. 176). Springer Nature.
9. Malecki, K. M., Keating, J. A., & Safdar, N. (2021). Crisis communication and public perception of COVID-19 risk in the era of social media. *Clinical infectious diseases*, 72(4), 697-702.
10. Matheus, R., Janssen, M., & Maheshwari, D. (2020). Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. *Government Information Quarterly*, 37(3), 101284.
11. Pangrazio, L., & Selwyn, N. (2019). 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. *New media & society*, 21(2), 419-437.
12. Pappas, I. O., Mikalef, P., Giannakos, M. N., Krogstie, J., & Lekakos, G. (2018). Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. *Information systems and e-business management*, 16(3), 479-491.
13. Reese, S. D., & Shoemaker, P. J. (2018). A media sociology for the networked public sphere: The hierarchy of influences model. In *Advances in foundational mass communication theories* (pp. 96-117). Routledge.
14. Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics—Challenges in topic discovery, data collection, and data preparation. *International journal of information management*, 39, 156-168.
15. Yanamala, A. K. Y., & Suryadevara, S. (2023). Advances in data protection and artificial intelligence: Trends and challenges. *International Journal of Advanced Engineering Technologies and Innovations*, 1(01), 294-319.
