FOCUS

Data news dissemination strategy for decision making using new media platform

Lu Li

Accepted: 18 January 2022 / Published online: 25 March 2022 © The Author(s) 2022

Abstract
In big data era, data has become an important part of human lives and work. At the same time, data plays an important role in information acquisition and dissemination. At present, the influence of data journalism is gradually increasing. However, unlike other countries, China data journalism started late. To study the problems of data journalism and the trend of future development, we use the method of combining data and news to explore the status quo and dig out the existing problems. This article first summarizes the research status, research methods, and theoretical basis of data journalism’s propagation path. Next, it uses Lasswell’s 5W model: a new model to analyze data news from five aspects, namely, disseminator, disseminating channel, dissemination content, audience, and dissemination effect. Finally, based on content analysis, searching, and data mining, an indicator system is constructed for the current new media’s news dissemination effect evaluation, and the Delphi method is used to assign weights to various indicators and make decisions based on them. By analyzing the results, this paper identify the problems in the process of combining data journalism and new media platforms, and provide help for the future communication strategy of data journalism.

Keywords Data journalism · New media platform · Decision making · Communication strategy · Communication effect · Lasswell’s 5W model

1 Introduction
In recent years, whether in our personal lives or work, the explosive growth of data has been surprising. Data is no longer an unfamiliar concept to us. It has been completely integrated into our lives and work. Both quantity and density have reached unimaginable levels. We are like a flat boat sailing in the vast ocean. We become inseparable from big data unconsciously and become “only in this mountain”. “Big data” is a vocabulary derived from the field of internet information technology, which refers to the massive amount of data generated in the internet information era and related technological innovations (Armstrong 2014). Big data has the characteristics of “4 V”, that is, massive data scale (volume), various data types (variety), extremely fast data flow (velocity), and low value density (value). What makes big data more prominent is that it emphasizes asking the right questions instead of just providing the right answers. With more data available, we need to re-examine existing assumptions and re-analyze existing data if we want to understand the world more deeply.

In this context, big data pays more attention and has a crucial role (Yousif 2015). Therefore, big data is naturally compatible with news, and the generation of data news is not new. Big data provides the news industry with massive amounts of data, allowing news practitioners to dig out the connections and relationships between various things in the data, thereby generating data news, a new type of news reporting method. In foreign countries, the British “Guardian” used data journalism for reporting as early as 2010; the US “New York Times” and the “Washington Post” also began to use data journalism as a new reporting method in 2014. News is truly accepted by the mainstream media. On the other hand, in China, some traditional media have also embarked on the road of data journalism and...
started to build data journalism studios, showing their due attention to data journalism. Relying on the technical support of the era of big data, journalism and communication have burst into new vitality in the new era. Just as the integration of communication and sociology, psychology, and other disciplines, the combination of data technology and news makes data journalism a new era. The latest achievements in data journalism are increasingly developing into a new subject. Research on data journalism can not only provide the latest academic achievements for the academic community, but also provide theoretical guidance and talent training directions for the news industry.

The following are the contributions of the proposed research:

- To summarize the research status, research methods, and theoretical basis of data journalism’s propagation path.
- To use Lasswell’s 5W model: a new media as a platform to analyze data news from five aspects, namely, disseminator, dissemination channel, dissemination content, audience, and dissemination effect.
- Finally, based on content analysis, searching, and data mining, an indicator system is constructed for the current new media’s news dissemination effect evaluation, and the Delphi method is used to assign weights to these indicators and make decisions based on these indicators.

The paper is structured as follows; Sect. 2 is the detailed representation of the related work concerning the proposed research. Research methods and theories are shown in Sect. 3. The description of experiments and discussions are given in Sect. 4. The paper is concluded in Sect. 5.

## 2 Related work

Data journalism originated from precision journalism. Philip Meyer’s series of reports on “People over Twelfth Street” in 1967 is considered by the academic circles to be the origin of accurate news reporting. Accurate news has cultivated the awareness of journalists to value data and laid a good foundation for future data-driven reporting. From the practical level of foreign data journalism, in 2009, data journalism entered the public eye. The British “Guardian” created a “Data Blog” column on its website. Since then, the “New York Daily”, “Wall Street Journal”, BBC and other news media have opened data news columns. The first International Data Journalism Roundtable hosted by Mirko Lorenz in 2010 opened the prelude to the research on data journalism in academia. In 2011, “The Data Journalism Handbook” (The Data Journalism Handbook) was published, which was the first book to systematically introduce data journalism. Judging from the academic research of foreign data journalism, it probably includes the following aspects: (1) Introduction to data journalism related technologies. Such as Veglis, Andreas and Bratsas, Charalampos put forward the importance of journalists mastering information and communication technology (ICT) skills in “Reporters in the Age of Data Journalism”, paying particular attention to Web 3.0 and open data that can play an important role in data journalism. (Veglis 2017).

Thomas Horky and Philipp Pelka’s (Data Visualisation in Sports Journalism), through five interviews with experts and an exploration of the content of three special blogs, analyzed the challenges faced by professional online sports news evaluation and data publishing (Horky and Pelka 2017). Simon Rogers’s “Data Journalism Trend: Unleash the Power of Visual Reporting” is published. The book not only comprehensively introduces how the author obtains data and uses data to explain news events, but also tells the production process of some classic cases of data journalism. And it used visualization to deeply interpret the social impact brought by data news. (2) Interpret data news through specific case analysis. Such as Tandoc and Edson C (Small Departures, Big Continuities Norms, Values and Routines in The Guardian’s Big Data Journalism) by analyzing the content of big data news reports in The Guardian, the pioneer of contemporary big data news, compared the big data the values, norms and practices of data journalism and traditional journalism (Tandoc 2017). Lim, Jeongsub’s (Representation of Data Journalism Practices in The South Korean and US Television News) investigated and studied how to use data journalism to represent television news by comparing and analyzing the data news content of Korean and the US television networks (Lim 2019). (3) The impact on the data journalism industry. (Borges-Rey 2016) Portilla, Idoia’s (Contributions of Academic Articles to the Practice of Journalism and Data Management) pointed out that more and more open data resources and the arrival of big data have promoted data as a news source, which is confirmed by SLR and content analysis. Data management in news practice requires not only skills in data visualization technology, but also skills in ethics, marketing or audience monitoring (Portilla 2018). (4) Analysis and research on data journalism education. For example, Bahareh R. Heravi’s (3WS of Data Journalism Education: What, Where and Who) pays particular attention to the education of data journalism in the global higher education sector (Heravi 2018). Bradshaw, Paul’s (Data Journalism Teaching, Fast and Slow) draws on the experience of teaching data journalism in various backgrounds over the past ten years, and analyzes the experience and
lessons of different teaching techniques and data journalism teaching choices (Bradshaw 2018).

With the development of computer technology, my country’s four major websites, Sohu, NetEase, Tencent, and Sina, have also launched data news production and launched their own data news columns since 2011, such as “The Way of Numbers” and “NetEase Digital Reading”, “News Encyclopedia” and “Illustrated World”. However, at the early stage of the development of data journalism, the media generally only presented news by making simple infographics, which was relatively simple and unattractive. With the gradual maturity of technology, Caixin.com established the Caixin Data Visualization Laboratory and launched the “Digital Theory”. In 2013, my country's academic circles began research on data journalism, and the number was relatively small. From 2014 to the present, the research literature of data journalism has grown rapidly, but data journalism is mostly analyzed or described from a theoretical level. As scholars gradually deepen their research on data journalism, research directions and perspectives have begun to show a diversified trend. Through analysis, it is concluded that domestic scholars’ research on data journalism is mainly carried out from the following aspects: analyzing the presentation form of data journalism by case. For example, Wen W H and Li B took the British “Guardian” as an example to analyze the implementation steps and characteristics of data journalism, and pointed out the challenges that data journalism poses to journalists (Wen and Li 2013). Discuss the issue of data journalism education. Peng L pointed out that in the era of media convergence, news media people need integrated thinking, the ability to cross between professional media and social media, and the ability to link content and products (Peng 2015). Shen H, Tan H, and Wen L studied the generation and development of data journalism by analyzing the evolution of news reporting methods, and on this basis, discussed how to carry out data journalism education (Shen et al. 2014). Analyze the theoretical concept and meaning of data journalism. It is represented by Yu G M’s book “The Big Data Era of News Communication”, which defines the definition of big data, describes the way of thinking of big data, and analyzes and demonstrates the application and dissemination of big data in the field of news communication. Through the integrated analysis of data journalism theory and practice, Lang J S and others have explored the innovative path of news visualization in the era of big data, such as news narration and interaction (Lang and Yang 2014).

Research status of data news dissemination path: The data news dissemination path refers to the distance and method of the dissemination experience of data news from the publisher to other audiences. Foreign research results on the transmission path of data journalism mainly focus on textual research or theoretical research. For example, to explore the future development path of news, Skye Doherty’s (Hypertext and Journalism Paths for Future Research) is based on computer science and the journalism literature, considering how hypertext can be used as a narrative tool, and specifically examines the narrative design and spatial visualization. Concepts and advance the research agenda around these topics (Doherty 2014). David Domingo, Pere Masip and Irene Costera Meijer (Tracing Digital News Networks towards an Integrated Framework of the Dynamics of News Production, Circulation and Use) distinguishes news practice from specific theoretical categories, overcomes the disciplinary gap between news production and news consumption analysis, and corrects problematic news norm principles (Domingo et al. 2015). The Mediated Data Model of Communication Flow: Big Data and Data Journalism by Veglis et al. tried to clarify the current communication process of data journalism by outlining the theory of analyzing communication models and using two-step communication theory as a starting point (Veglis and Maniou 2018).

3 Research methods and theories

3.1 Research methods

The research method of this work is mainly based on the literature research method, supplemented by the case research method and content analysis method. After selecting data journalism as the research object, I searched for different journals, monographs, and papers in the literature by using “data journalism,” “communication strategy,” and so on as the key words, so as to be aware of it. This article mainly uses different data news columns in the new media platform, including Guyu Data, NetEase “Digital Reading,” Caixin Data Theory, etc., as research cases to study the style, topic selection, and visualization characteristics of different data news columns. The same news or the same type of data news reports in different columns are selected as the research samples, and their column classification, news content, and news reading volume, number of comments, reposts, etc., are used as indicators to provide data support and reference for the research of this article.

3.2 Data journalism concepts and relationships

Data journalism is also called “data-driven journalism”. It is a new way of combining news and data in the era of big data. Open data, rigorous narrative logic, and appropriate visual presentation form the basis of data journalism. In 2006, Adrian Harowa, the founder of EveryBlock, first
proposed that journalists can abandon the traditional way of reporting in words by publishing data. The American press believes that data journalism is different from computer-assisted reporting. They believe that computer-assisted reporting is mainly a means of computer-assisted news reporting, while the focus of data journalism is on data. In China, data visualization and information charts are often accompanied by data news. They are often unclear by people “Li K and Li G”, but they are fundamentally different. Today’s data news is not just about static data. In some major events and events, it is accompanied by the dynamic output of data. For example, the changes of medals in the Olympic Games present the advantage of being easy to read and understand in the visual display of data news. Finally, it becomes a tool for dynamic reporting.

3.3 5W theory

The American political scientist Harold Laswell first proposed the five basic elements of the communication process in a paper entitled “The Structure and Function of Communication in Society” in 1948. They were arranged to form a process model later called the “Five W Model” or “Lasville Formula”. The five W are the first letters of the five interrogative pronouns in English, namely: Who; Says What; In Which Channel; To Whom; With What Effect. A piece of data news also includes these elements in the dissemination process, forming a complete dissemination process. In the new media platform, the 5W factor may not seem obvious, but it still becomes an important factor affecting the dissemination of data news (Yirmiya 2010).

3.4 Delphi method weight analysis

In the ever-changing new media era, accurate and effective evaluation of communication effects can have greater significance in the design, implementation and decision making of communication activities. Through the judgment of the corresponding indicators, the author will use the Delphi method to construct the weight of the communication effect evaluation index system based on the various indicators in the news communication. The indicator data are all from the ADNI1 data set.

3.5 The main process of the Delphi method

(1) Formulate evaluation events: indicators for the evaluation of news dissemination effects in the new media environment, clarify the evaluation goals, compile evaluation event tables, analyze and process the corresponding indicators, determine the prediction theme according to the nature and content to be evaluated, and select the corresponding expert group responded to opinions based on the needs of the research content. (2) According to the requirements of the forecast theme, the content of the survey and the purpose of the survey, through the collection and sorting of data, several questions were raised and the questionnaire was designed. (3) Implement answers to inquiries and conduct multiple rounds of feedback. In the first round of questionnaires, the background and research contents of the above-mentioned topics are sent to the surveyed experts, and the experts will make corresponding written answers after analyzing and researching the questions in the questionnaire. After collecting the forms, summarizing and sorting out the second round of questions and problems that need improvement, the second round of questionnaire is sent to the experts who were surveyed in the previous round for further feedback. (4) The process of collecting opinions lasts for three rounds, until all experts no longer change their opinions. (5) Carry out necessary data analysis and processing on the final results, and draw conclusions of evaluation and prediction.

3.6 Numerical processing

(1) Calculate the arithmetic average of each expert’s score for each object.

\[ M_i = \frac{1}{m} \sum_{k=1}^{m} C_{ik} \]  

where \( M_i \) represents the arithmetic average of all experts’ scores on the evaluation index \( i \) object (a certain index index); \( m \) represents the number of experts participating in the evaluation of the \( i \) index; \( C_{ik} \) represents the score of the \( k \) experts on the index \( i \). When the researcher adopts the tenth scale, the arithmetic mean value \( M_i \) is 1–10. The greater the value of \( M_i \), the greater the relative importance of the object among all the objects participating in the scoring.

(2) For each round of scoring by experts, it is also necessary to calculate the standard deviation and coefficient of variation of the scoring object. Calculate the standard deviation of a certain scoring object.

\[ \sigma_i = \sqrt{\frac{1}{m} \sum_{k=1}^{m} (C_{ik} - M_i)^2} \]  

where \( \sigma_i \) is the standard deviation of the scoring index \( i \); \( C_{ik} \) represents the scoring of index \( i \) by \( k \) experts; \( M_i \) represents the arithmetic mean of all experts’ scoring of index \( i \), and \( m \) represents the number of experts participating in the evaluation of
index $i$. The standard deviation represents the degree of variation in the evaluation. The smaller the standard deviation, the more consistent the assignment of indicator weights by experts. Calculate the coefficient of variation:

$$
V_i = \frac{\sigma_i}{M_i}
$$

where $V_i$ represents the degree of coordination of the opinions of the experts on the $i$-th score index. The smaller the value of $V_i$, the greater the degree of coordination of the experts high. At this time, the experts' feedback on the index is gradually becoming similar. The more the opinions are closer to the index, the greater the reference value of the calculated weight. After scoring the four dimensions of the above-mentioned news release indicators, and optimizing the weights of the indicators based on the historical model, a more scientific judgment can be made on the corresponding news dissemination effect.

4 Experiments and results

4.1 Topic selection preference and distribution of new media platforms

Through the research of the previous literature and the reading of the big data news columns, the author roughly divides the topics of data news into eight categories: international, political, economic, cultural, social, environmental, popular science, and others (including sports, entertainment). After statistically sorting out all the data news works of the two media NetEase “Study Reading” and People’s Daily‘’ Illustrated News’’ for the whole year of 2019, the results are as follows:

As shown in Fig. 1, the statistical results of the WeChat public account People’s Network Graphic News show that the distribution of selected data news works published by Graphic News includes 14 international articles, 55 political articles, 8 economic articles, 14 cultural articles, and social articles., 6 articles, 2 environmental articles, 3 popular science articles, and 9 other articles. Compared with the report on the economic aspects of “Digital”, the content of “Illustrated News” of People’s Daily Online focuses on the political field. The proportion of topics in the political field is as high as 53%, which matches the positioning of People’s Daily Online. Pay attention to news reports on political activities, political events, major conferences, and leaders’ visits at home and abroad.

The statistical results of the topic selection distribution on NetEase Weibo show that the topic selection includes 1 international, 1 political, 4 economics, 12 cultural, 40 social, 3 environmental, and popular science, 5 articles and 10 other articles. From this we can see that NetEase’s original data news topics are more inclined to social and cultural fields. These two topics account for 75% of the total topics. The questions are very diverse, including regional customs, music culture, ancient and modern history, food exploration, north–south differences, etc., such as “How do Chinese love hot pot”, “China’s obesity map is released, and northerners have thicker waists”. These reports, through the collection, analysis and summary of data, give readers a new understanding of the current culture in various aspects of our country. However, the topic selection is not timely and newsworthy, but it is full of fun (Fig. 2).

4.2 The dissemination channels of data news on new media platforms

In the new media era, WeChat, Weibo, and clients have become new platforms and battlefields for the mobile Internet. Nowadays, the information reception of most netizens is not affected by the “two micro-ends”. The new media platform is also a new communication channel for the development of data news in the future. It can not only increase the readers of data news, but also contribute to data news. Communication plays an important role. The morphological indicators of news dissemination channels are shown in Tables 1 and 2; the comparison between the channels is shown in Figs. 3 and 4.

4.3 Audience’s use and satisfaction of data news

According to media impressions, people choose a specific medium or content to start a specific contact behavior; there are two results of contact behavior, one is satisfaction, the other is unsatisfied; whether satisfied or not, this result will affect future media contact behavior, people will modify the existing media impressions according to the results of satisfaction, and change their expectations of the media to varying degrees (Wu 2010). Compared with traditional news, data news has a stronger visual impact, which will deepen readers’ impression of news and achieve a more efficient form of communication. The basis of data journalism is data. The use of data as the basic element to tell news can make an in-depth connection between data, news, and visualization, which can help readers understand the complex news content and background. This is also data journalism. This advantage allows readers to replace some lengthy traditional news by reading short data news, making readers more inclined to choose data news as a way
to receive information. However, the amount of information in data news itself is much larger than that of other types of news, so the more important thing is how to present the data in a reasonable way. In other words, readers will pay more attention to the presentation method. Table 3 shows the index weight of each news form. Figure 5 shows a comparison chart.

4.4 The problem of data news dissemination based on new media platforms

(1) Data journalism team building is lagging behind, and there is still a significant gap between domestic data journalism team building and foreign countries. When domestic data journalism was only used as a branch of online journalism, foreign data journalism teams had already begun commercial operations, and even completed a complete industrial chain, and commercial
crowdsourcing appeared. Although domestic data journalism studios were established late, this is not the reason or excuse for the lagging development of data journalism in China. The contempt for data journalism is the main reason for restricting the development of data journalism (Du 2019). (2) The dissemination of data news lacks innovation, and the content of data news is its top priority. Although the expression of data news is very different from traditional news, there are still many problems in its content, including the abuse of visual forms, reporting the content lacks depth, and problems such as patching together data need to be solved urgently. (3) The data news dissemination path is solidified and single. In the analysis of the dissemination path of the data news new media platform, it is found that the dissemination path of data news in the new media platform is relatively solidified, and the data news delivery rate and the secondary transmission
rate even appear obvious. Contrast (Watters et al. 1998). (4) Data news is difficult to attract audiences. Technology promotes media progress. New media platforms such as new media have emerged. The emergence of new media platforms has also given each audience a “microphone” that can speak out. This is technology to the audience. From the original one-centered communication model to today’s decentralized communication, the audience is in a divergent situation, and its dependence on a single medium is greatly reduced. For example, in the rise of short videos, today’s audiences some “two micro-ends” shifted to Douyin and Kuaishou, which distracted their media usage time. This means that data journalism still has strong competitors, and its opponents may not be journalists. If data journalism is stagnant and changes are made, it will be difficult to attract an audience.

5 Conclusion

This paper examines the current state of data news dissemination on the new media platform from a variety of perspectives, i.e., the literature reviews, a macro-study of the data news industry and team, and a micro-study of individual data news cases. It include communication subject, communication content, communication channel, communication audience, and communication effect. This paper gives a detailed explanation of the current development of data journalism on the current new media platform and puts forward the corresponding data dissemination strategy, hoping to provide suggestions for the development of the industry and also provide some value for academic research. The innovation of this article lies in the designing contents from various perspective and making decisions based on Lasswell’s 5W model and Delphi models. It provides a unique perspective for the researchers with new directions, as well as detailed and rich content on various media outlets. The decision making capabilities of Lasswell’s 5W model in combination with Delphi model provides results that are convincing and better than the existing approaches. In the future, the proposed idea can be implemented using models other than Lasswell’s 5W and Delphi models to see the results on various media and news outlets.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. The paper did not receive any financial support.

Data availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The author declared that I have no conflicts of interest to this work.

Ethical approval The paper does not deal with any ethical problems.

Informed consent The author declare that I have informed consent.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Armstrong K (2014) Big data: a revolution that will transform how we live, work, and think. Math Comput Educ 47(10):181–183
Borges-Rey E (2016) Unraveling data journalism. Journal Pract 10(7):833–843
Bradshaw P (2018) Data journalism teaching, fast and slow. Asia Pac Med Educ 28(1):55–66
Doherty S (2014) Hypertext and journalism paths for future research. Digit Journal 2(2):124–139
Domingo D, Masip P, Costera Meijer I (2015) Tracing digital news networks towards an integrated framework of the dynamics of news production, circulation and use. Digit Journal 3(1):53–67
Du R (2019) Dancing with chains: a case study of data journalism in China. Int J Inf Syst 12(2):225–268
Heravi BR (2018) 3WS of data journalism education: what, where and who. Published online 18:349–366
Horky T, Pelka P (2017) Data Visualisation In Sports Journalism. Digit Journal 5(5):587–606
Lang JS, Yang H (2014) Data journalism: innovative path of visualized news dissemination in the era of big data. J Commun Univ China 36(03):32–36
Lim J (2019) Representation of data journalism practices in the South Korean and US television news. Int Commun Gaz 81(1):89–113
Peng L (2015) In the era of convergence, where is new media education heading. News Writ 03:5–7
Portilla I (2018) Contributions of academic articles to the practice of journalism and data management. Observatorio (OBS) 12(3):66–82
Shen H, Tan H, Wen L (2014) “Data journalism” development and “data journalism” education. J Commun Univ China 36(11):139–142
Tandoc EC (2017) Small departures, big continuities norms, values and routines in the guardian’s big data journalism. Journal Stud 18(8):997–1015
Veglis A (2017) Reporters in the age of data journalism. J Appl Journalism Med Stud 6(2):225–244
Veglis A, Maniou TA (2018) The mediated data model of communication flow: big data and data journalism. Kome-an Int J Pure Commun Inq 6(2):32–43
Watters CR, Shepherd MA, Burkowski FJ (1998) Electronic news delivery project. J Am Soc Inf Sci 49:134–150

Wen WH, Li B (2013) Data news reporting in the era of big data — taking the British “Guardian” as an example. J Commun Univ China 35(05):139–142

Wu JZ (2010) Analysis of the problems and countermeasures in the current promotion of mobile TV. News Knowl 05:35–37

Yirmiya N (2010) Editorial: early prevention and intervention—the five W (and one H) questions. J Child Psychol Psychiatry 51(12):1297–1299

Yousif M (2015) The rise of data capital. IEEE Cloud Comput 2(2):4–4

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.