News Industry under the Background of Big Data

Yixin He

1Macau University of Science and Technology, Macao, China

Abstract. With the continuous update of technical means, it is possible to store and process a large amount of data in news dissemination. This research mainly discusses the development of the news industry in the context of big data. Entering the era of big data, with the update of data collection and analysis and processing technology, audience feedback in news dissemination can directly through the in-depth mining of audience data and full sample analysis, fully and timely present the audience's needs, preferences and behavior patterns. The value of audience feedback. Big data analysis of users can not only be applied to the overall analysis of users, but also can be refined to individual analysis. Through big data-based charts, the originally singular information presents different expressions, which greatly enhances the readability of news information. The main factors that affect users' use of mobile news clients are the speed of news updates and the professionalism and comprehensiveness of news content, which account for 30% and 20% respectively. This research helps to propose a feasible development path for the future development of traditional news production.

1 Introduction

In the news dissemination industry, the exponential growth of big data and the application of related technologies also pose many challenges to traditional news production. Not only has the professional status of traditional journalists been challenged as never before, but in the way of news production, big data has overturned traditional news production in the three aspects of news information collection, processing and distribution.

Among media organizations, to complete the production of a piece of news information, the first thing that should be done is to take various means to collect relevant news-worthy information. This is the first and most important step in the news production process. One step [1-2]. In our daily life, various events happen all the time, and the collection of news information is to select news-worthy information from these massive events and activities to provide raw materials for the next news production[3 -4]. In terms of objectivity, traditional news production relies on news sensitivity to discover news, which leaves most of the decision-making power of news discovery to news practitioners, although most professional news practitioners can rely on professional accumulation and practice to complete news. Information collection work [5-6]. But because news sensitivity is a kind of thinking that exists in the minds of journalists, it will inevitably be affected by the subjective ideas of journalists, thereby affecting the objectivity of news information [7-8]. The news discovered through user big data analysis has absolute objectivity. It is completely obtained through the analysis of massive objective data. These data are generated based on the user's natural activities and behaviors, without any human factors or Subjective concepts, data reflected are objective facts [9-10].

With the continuous updating of technical means, it has become possible to store and process a large amount of data in news dissemination, which promotes the development of news dissemination and sentiment. More and more news dissemination agencies are involved in the development of public opinion monitoring and analysis platforms. and service. The effects and feedback of news dissemination have become one of the focuses of attention. Accurate data enables news media to better adjust reporting strategies and methods. Public opinion service has also become a brand business.

2 Development of the News Industry

2.1 News Industry

The data news of these media that have carried out the practice of data news accounts for a small proportion of the media's total news production. Basically, they open a data news column. Currently, no media completely takes data news as the mainstream news form. Basically, they are in the stage of pioneering and innovative new things. The difference between different media is that some media are early, fast, and the standard of existing data news production is relatively high, while some media are still in the bud and learning state. While accepting the era of big data, the news industry and academia clearly discovered the changes that data has made to news. From the perspective of news carriers and audience acceptance,
2.2 Big Data

What data journalism gives to news dissemination is a new data thinking and data concept, and what it brings is a new mode of news discovery and report presentation. The requirement of data journalism is that news communicators must master the ability of mining and observation and analysis of news data, based on the assistance and support of new technologies, presenting a deeper level. The second and more valuable news information, which is also data journalism and traditional news. The essential difference in reporting mode.

The space overhead of the division process is \( k \) and the data storage space is \( n \), so the space complexity is:

\[
S(n) = m + k + n
\]  

(1)

The evaluation \( Q_{ov} \) function of the divided community can evaluate the structure of big data.

\[
Q_{ov} = \frac{1}{m} \sum_{i=1}^{m} \left[ \sum_{j=1}^{n} F(\alpha_{i,c}, \alpha_{j,c}) A_{ij} - \frac{F_{\text{out}}^{\text{out}}}{m} \right]
\]  

(2)

\[
F(\alpha_{i,c}, \alpha_{j,c}) = \frac{\sum_{j} F(\alpha_{i,c} - \alpha_{j,c})}{(1 + e^{-f_{\alpha_{i,c}}})(1 + e^{-f_{\alpha_{j,c}}})}
\]  

(3)

Among them, \( A \) is the network proximity matrix. Big data makes the potential connection of data between industries appear. In the era of "small data", most of the information storage and communication are still stored as paper documents. In the era of big data, due to the ubiquity of data, data is the main method of storage and communication. The improved technology can decrypt these data through the application program to discover its inner connection.

3 Exploring Experiments in the News Industry

3.1 User-Centric

With the diversification of news information acquisition channels and methods, news dissemination has changed from a transmitter-centered to a receiver-centered. "User Center" is embodied in the field of news dissemination. The most notable feature is the emergence of personalized news demand, that is, the systematic mining, screening, classification and integration of news data, and the use of multiple media and formats to meet the needs of news users. There is a growing demand for segmented and personalized news selection. Entering the era of big data, with the update of data collection and analysis and processing technology, audience feedback in news dissemination can directly through the in-depth mining of audience data and full sample analysis, fully and timely present the audience’s needs, preferences and behavior patterns. The value of audience feedback.

3.2 Enhance the Value of User Analysis

Big data analysis of users can not only be applied to the overall analysis of users, but also can be refined to individual analysis. Through big data-based charts, the originally singular information presents different expressions, which greatly enhances the readability of news information. Table 1 shows the establishment of data news columns of some Internet news portals and traditional media new media portals in China.

| Media name        | Data news column name | Media name        | Data news column name |
|-------------------|-----------------------|-------------------|-----------------------|
| Xinhuanel         | Data news             | Sina              | Illustrate the world  |
| People's Daily    | Graphic news          | Caixin            | Data theory           |
| Online            |                       |                   |                       |
| Sohu              | Data Way              | Nandhuaswiyang    | Southern Metropolis   |

4 Development of the News Industry in the Context of Big Data

4.1 News Production Cycle Analysis

As far as the news dissemination industry is concerned, professional data technology should be used in news production to collect, clean, analyze and organize information on specific topics. Finally, it is necessary to visualize news products on the end of electronic equipment. A few simple the steps require huge manpower, financial resources, and energy. This requires coordination and cooperation between different departments within the news organization. Different departments are responsible for their respective areas of expertise, and ultimately maximize production efficiency, even in some cases. Cross-border cooperation is required. Therefore, journalists in the new era should change their own-centered production philosophy and replace them with team-integrated and cooperative production concepts. It can be seen from Table 2 that during the production of general data news data, 18 people produced an average of 1.72 days per person in January, and the overall production frequency was 19 times.

| Factor                  | Quantity | Average production time (days) | Total production times | Average number of people involved in production |
|-------------------------|----------|--------------------------------|------------------------|-----------------------------------------------|
| Time                    |          |                                |                        |                                               |
| January (31 days)       | 18       | 1.72                           | 19                     | 1.05                                          |
| February (29 days)      | 27       | 1.07                           | 30                     | 1.11                                          |
| days) | Total | 45 | 1.11 | 49 | 1.08 |

### 4.2 Convenience of Information Transmission

The convenience analysis of information transmission is shown in Figure 1. Different from the traditional news media, the aggregate news client is carried on the mobile device, and the user receives the news transmission process is very fast and convenient, without having to buy newspapers or log in to the portal as in the past. The main factors affecting users' use of mobile news clients in 2018 are the speed of news updates and the professionalism and comprehensiveness of news content, which accounted for 30% and 20% respectively. Interactive comment function, beautiful interface and accuracy of customized content accounted for 15%, 13% and 12% respectively. In addition, advertising accounts for 10%. It can be seen that in the era of information explosion, users are more pursuing the update speed of news, paying attention to the quality and professionalism of news content. Aggregated news clients can provide fast-updated news information, which greatly improves user experience. Based on big data analysis, the aggregated news client can realize larger-scale news dissemination, and the content coverage is wider. Information search and information push are very convenient. It breaks through the limitations of traditional media manpower and environment, and has a rich variety of information. Data to find their reading interest, and finally realize the convenient and dynamic transmission of news information.

![Figure 1. Convenience analysis of information transmission](image)

### 4.3 Quantity Analysis of News Data

The statistics of the number of data appearing in all sample articles in 2019 were calculated, and the following results were obtained: Sina's "Illustration World" averaged 13.5 data per article; Sohu "Digital Way" had an average of 18.3 data per article; NetEase "Numerical Reading" averages 28.7 data per article. The Guardian has an average of 26.4 data per article; the New York Times has an average of 23.3 data; ProPublica has an average of 198.3 data per article. Put these data back into the original application, and then look at the comparison of the data, you can summarize the characteristics of the two layers. First of all, the amount of data basically determines the direction of the article and the length of the article. The less the amount, the less content of the news. The greater the number, the greater the readability of the news, the greater the amount of information, and the greater the value of news. Secondly, the data in domestic data news has some functions that are reduced to embellishment, serving the main line of news. The data in foreign data news plays a leading role and determines the main line and direction of news. The results of the analysis of the number of news data are shown in Figure 2.

![Figure 2. Quantity analysis of news data](image)
5 Conclusion

With the diversification of news information acquisition channels and methods, news dissemination has changed from a transmitter-centered to a receiver-centered. "User Center" is embodied in the field of news dissemination. The most notable feature is the emergence of personalized news demand, that is, the systematic mining, screening, classification and integration of news data, and the use of multiple media and formats to meet the needs of news users. There is a growing demand for segmented and personalized news selection.

Entering the era of big data, with the update of data collection and analysis and processing technology, audience feedback in news dissemination can directly through the in-depth mining of audience data and full sample analysis, fully and timely present the audience’s needs, preferences and behavior patterns. The value of audience feedback. This research helps to propose a feasible development path for the future development of traditional news production.

References

1. Yudong C, Yuejie C. Harnessing Structures in Big Data via Guaranteed Low-Rank Matrix Estimation[J]. IEEE Signal Processing Magazine, 2018, 35(4):14-31.

2. Al-Ali A R, Zualkernan I A, Rashid M, et al. A smart home energy management system using IoT and big data analytics approach[J]. IEEE Transactions on Consumer Electronics, 2018, 63(4):426-434.

3. Mohammadi M, Al-Fuqaha A. Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges[J]. IEEE Communications Magazine, 2018, 56(2):94-101.

4. Duan M, Li K, Liao X, et al. A Parallel Multiclassification Algorithm for Big Data Using an Extreme Learning Machine[J]. IEEE Transactions on Neural Networks & Learning Systems, 2018, 29(6):2337-2351.

5. Mann L. Left to Other Peoples' Devices? A Political Economy Perspective on the Big Data Revolution in Development[J]. Development and Change, 2018, 49(1):3-36.

6. Andre J C, Antoniu G, Asch M, et al. Big Data and Extreme-Scale Computing: Pathways to Convergence - Toward a Shaping Strategy for a Future Software and Data Ecosystem for Scientific Inquiry[J]. International Journal of High Performance Computing Applications, 2018, 32(4):435-479.

7. Neggers J, Allix O, Hild F, et al. Big Data in Experimental Mechanics and Model Order Reduction: Today's Challenges and Tomorrow's Opportunities[J]. Archives of Computational Methods in Engineering, 2018, 25(1):143-164.

8. Prasad S, Zakaria R, Altay N. Big data in humanitarian supply chain networks: a resource dependence perspective[J]. Annals of Operations Research, 2018, 270(1):383-413.

9. Samiya Khan, Kashish Ara Shakil, Mansaf Alam. Cloud based Big Data Analytics: A Survey of Current Research and Future Directions[J]. Computer Science, 2018, 03(5):107-117.

10. Yang L T, Kuang L, Chen J, et al. A Holistic Approach to Distributed Dimensionality Reduction of Big Data[J]. IEEE Transactions on Cloud Computing, 2018, 6(99):506-518.