Interpretation of Big Data in Communication

Caixia Zhang[1]
School of humanities and law
Wuhan Donghu University
Wuhan, China
19689513@qq.com

Bo Zhang[2,*]
School of Computer Science
Wuhan Donghu University
Wuhan, China
*Corresponding author's
e-mail: bob.cheung@ovspark.com

***Abstract*: "Big data" has changed from a technical buzzword into a social wave, which has impacted the fields of information, politics, business, etc. As the producer and owner of data, the media industry is also affected by big data. The impact of big data on news communication is mainly reflected in the following aspects: enhance the communication effect, better meet the needs of the audience and predict the behavior of the audience; beneficial to the content innovation and strengthens the objectivity and profundity of news; enhance audience interaction and further enhance the value of audience feedback; promote the integration of media talents and put forward higher requirements for media literacy.**

Keywords: big data; communication; communication effect

I. INTRODUCTION TO BIG DATA

The New York times has declared that the age of big data has arrived in 2012.Since then, big data has quickly become a global topic, and from a technical buzzword to a social wave. The media industry, as the producer and owner of data, is also affected by big data. This paper will interpret big data from the perspective of the development of the media industry and provide a reference for the industry.

The industry generally summarizes the characteristics of big data into four "V": Volume, Variety, Velocity, Value. Big data first emphasizes the big data. According to liu jianming of tsinghua university, "big data is another name for the vast amount of information in vast amounts of data, a synonym for the information explosion that prevailed in the late 1980s."1The volume of big data is huge. The initial unit of measurement is not G, but P (1000 T), E (1 million T), or Z (1 billion T). This data includes web logs, video pictures, location information, and so on. Second, big data emphasizes value information. Massive data contains a lot of junk information, and the data value density is low. The truly valuable information is only a part, or even a very small part. How to quickly and timely analyze and obtain valuable information in a large number of data is the core and key of big data in this age of information explosion.

The world produces a huge amount of data every year. How to face and deal with these massive data is particularly important. The three major Internet companies, Tencent, Alibaba and Baidu are all actively promoting big data application systems. The big data engine launched by Baidu has multiple functions such as big data analysis, big data forecasting, and big data marketing. On December 12, 2019, according to data released by Baidu at the "2019 Baidu Knowledge Summit": Baidu's daily search volume of knowledge content has reached 1.54 billion times, and Baidu's knowledge-based products have exceeded 230 million daily users. According to the big data of education search released by Baidu, the Internet has become the main channel for people to acquire knowledge. From February to March, an average of more than 350 million people search and browse education-related knowledge on Baidu every day, up 86% year-on-year. The average daily video playback volume is 150 million.

II. THE IMPACT OF BIG DATA ON NEWS COMMUNICATION

In 2011, the world-renowned consulting firm McKinsey pointed out in the research report of "Big Data: The Next Frontier of Innovation, Competition and Productivity" that "the era of big data has arrived", and data has penetrated into every industry and business function area, and gradually became Important production factor. Faced with the impact of big data technology, some people in the media industry have proposed that "data has become the new core competitiveness of enterprises and countries"2.How does big data affect news communications? How should we look at big data? This is the question this article will explore.

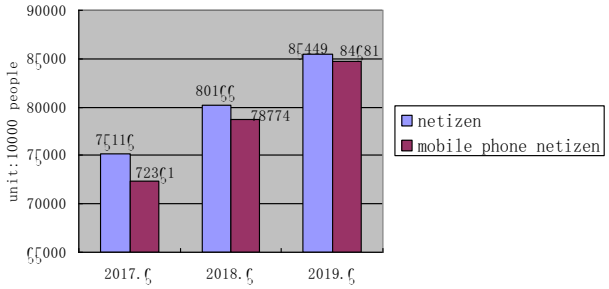
A. Enhance communication effects to better meet audience needs and predict audience behavior.

The advantage of big data technology is predictive possibilities. In the face of complex and diverse data, big data can find correlations through cross-analysis and predict various possibilities. The big data engine launched by Baidu has the functions of predicting economic indexes, scenic spot comfort, and disease prediction, which fully takes advantage of big data prediction. In news communication activities, the analysis of big data can better identify the audience, understand the characteristics of the audience, and then improve the effectiveness of news communication. According to statistics from the China Internet Network Information Center, as of June 2019, the scale of China's Internet users has reached 854 million, and the number of mobile phone users has reached 847 million (Table1). The proportion of Internet users using mobile

978-1-7281-5904-1/20/$31.00 ©2020 IEEE
DOI 10.1109/ICCEA50009.2020.00021

phones to access the Internet has reached 99.1%. A lot of useful information is hidden behind data such as comments, search, and viewing behaviors of Internet users. Using big data technology can summarize their behavior habits, personal preferences, and psychological tendencies, and then predict their viewing behaviors. Through a large amount of user data analysis, Netflix in the United States found an interesting phenomenon: users who love to watch the "House of Cards" in 1990, also like director David Finch, and also like Oscar actor Kevin Spacey. Therefore, the new version of "House of Cards" chooses directors, actors, scripts, and plots according to the audience's preferences, which is very popular after broadcast. The hit of "House of Cards" is not accidental, but a successful application of big data. After mastering the characteristics of the audience, it is a successful work to predict the audience's viewing behavior according to the needs of the audience. The use of big data technology can analyze the browsing data of netizens, and then recommend appropriate news to them based on their preferences and needs, to achieve accurate communication, improve the effectiveness of communication, and attract more news audiences.

Table 1. 2017-2019 Netizen scale chart



B. *Conducive to content innovation and strengthen the objectivity and depth of news.*

Big data is a quantitative tool that excels at speaking with data. In the traditional journalism era, data is also used in news reports, but mainly to assist news reports. With the explosion of information and the popularization of social media, data journalism has become a form of news. Different from the traditional text-based news, data news is data-first, text-first, emphasizing the deep mining and analysis of data. Data journalism is based on public data collection and analysis, which is conducive to discovering the deep-seated problems behind the news phenomenon, improving the media's grasp of social issues, and strengthening the objectivity and depth of news to a certain extent. During the riots in London in 2011, the British "The Guardian" used data journalism to conduct an in-depth analysis of the progress of the incident and the reasons behind it, and concluded that "Twitter has more advantages than disadvantages in the riot" 3, pointing out that Twitter is not just spreading rumors , also played a role in correcting rumors and mobilizing people to restore neighborhood order. This has clarified the rumors to a certain extent, negating the government's plan to shut down social networking sites such as Facebook and Twitter. The use of big data in news reporting has improved the objectivity and depth of news, expanded the influence of news, and received unanimous praise from the industry.

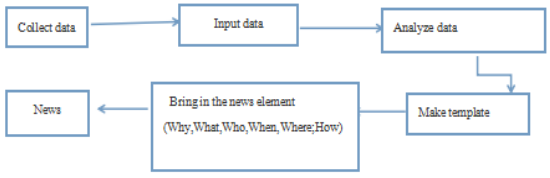
C. *Strengthen audience interaction and further enhance the value of audience feedback*

In the radio and television era, audience feedback is mainly through readers' letters and calls, and the timeliness and breadth of audience participation have certain limitations. Journalists are unable to accurately grasp the audience in the process of producing programs, do not understand what the audience is, and do not know the audience's behavior habits, emotional characteristics, and attitudes. Under the technical background of big data, the interactivity of communication activities is enhanced, and the feedback value of the audience is further enhanced. First of all, the Internet behavior of Internet users is part of the communication activities. The media contact behavior of netizens, comments, and other information will become data sources, and the recipients will collect and refine them to capture valuable information. For example, when analyzing the "Evening News" program, it is found that most of the audience are men. By tracking the traces of the public accounts and Taobao shopping, they can basically judge their personal characteristics, such as income status, media contact habits, hobbies Etc. On this basis, targeted advertising can be provided. Secondly, the feedback of the audience can directly determine the content and direction of the communication, affecting the effect of the communication. In the case of the hit TV series "House of Cards" in the United States, netizens can not only watch the program content in real time through the Internet, but also provide feedback through the interactive platform. The producers arrange the plot direction, scene content, plot details, and mechanism based on the opinions of netizens. Bit settings, etc. It can be seen that big data can play a role in strengthening audience interaction and further enhancing audience feedback value.

D. *Promote the integration of media talents and put forward higher requirements for media literacy*

At present, media at home and abroad have begun to use robots to write scripts. The writing robots include WordSmith of the Associated Press, Heliograf of the Washington Post, Blossom of the New York Times, Xinhua News Agency's quick pen Xiaoxin and Tencent's dreamwriter. Based on big data processing technology. First collect data, input the data into the database, then analyze and process the data according to the frequency of sentence occurrences and keywords of news elements to produce a set of templates that conform to the media posting style, and then insert the news element 5W1H into it, a news message produced. (Table 2)

Table2:Robot writing process



The new mode of computer writing press releases will pose a threat to the employment of journalists, a phenomenon that has attracted the attention of scholars. The industry and academia have also been discussing the adjustment of the media talent structure. Jianhua Zhu of the City University of Hong Kong once stated, "The news organization should include research groups, design groups, and publishing groups. The research group should have data search, data mining, and statistical analysis capabilities. The design group should have text writing and analysis capabilities. They should have the ability to publish information on online platforms and traditional platforms. Journalists with these capabilities can be regarded as qualified journalists. "4This puts forward higher requirements for media professionals' professionalism. to understand professional knowledge and data technology, it is necessary to increase the training of media personnel's data ability.

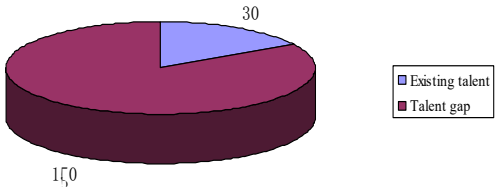
III. Countermeasures for the Development of Media in the Background of Big Data

A. Strengthen the training of big data technology professionals.

Big data can expand the depth of news, enhance the objectivity of news, and also enrich the expression of news and improve the effectiveness of news dissemination. But most journalists lack data thinking and have limited data analysis capabilities, which limits the use of data. Therefore, it is necessary to increase the training of news media personnel and improve their data awareness and data application capabilities.

Increase the training of data technology professionals. Data collection, analysis, and processing based on big data information technology require not only the participation of news and media talents, but also the technical support of computer programmers, data analysts, and other professionals to deepen the data to achieve the goals and transform them into Recognizable text and graphics. However, China currently has a big gap in big data talent. The "Digital Transformation of the Chinese Economy: Talents and Employment" report released by Tsinghua University shows that China's big data talent gap is 1.5 million and will reach 2 million by 2025 (table 3). It can be said that the shortage of big data talents is an important factor restricting the development of big data. In this context, the journalism and communication industry should increase efforts to train professionals in big data technology, and then promote the development of the media industry in accordance with the trend.

Table3:Big data talent gap chart
(unit:10000 people)



B. Pay attention to the integration and development of big data and artificial intelligence.

In recent years, China has successively released important guiding documents such as the Outline of Action for Promoting the Development of Big Data and the New Generation Artificial Intelligence Development Plan, which is committed to promoting the development of big data and artificial intelligence. Practice has proved that technology-based, algorithm-based big data The connection between artificial intelligence and big data is becoming closer and closer, and the integrated development of big data and artificial intelligence is complementary.

Baidu actively grasps the historical opportunity of the integration of big data and artificial intelligence, and has achieved a leading position in the development process of artificial intelligence and big data fusion. Based on the massive data of the search platform, Baidu has expanded its business scope to the integration of artificial intelligence and big data., Cloud computing and vertical industry applications (mainly industrial, medical, government, education, media, etc.) and other fields.

According to the big data of education search released by Baidu, the Internet has become the main channel for people to acquire knowledge. From February to March, an average of more than 350 million people search and browse education-related knowledge on Baidu every day, up 86% year-on-year. The average daily video playback volume is 150 million. During the pneumonia epidemic prevention, "online lessons" became an important way for students to learn. In the past 30 days, the search for online courses has skyrocketed, an increase of 1891% month-on-month. "What to do if the online course suddenly gets stuck" and "Will the online screen split screen teacher find out" have become hot issues. By analyzing these data, we can understand the needs and opinions of the public.

As the leading search engine platform in China, Baidu successively released the "Baidu Health Online Medical Big Data Report" and "Baidu 3.15 Search Big Data Report 2020". The report shows that the Internet medical industry experienced explosive growth during the epidemic, and Baidu Health Search enthusiasm of online consulting platforms such as ask a doctor has increased by 96%, and has provided users with more than 25 million services in total, reflecting the high demand for online consultation during the epidemic prevention period. By analyzing these data, we can understand the needs and opinions of the public.

IV. Conclusion

The British scholars Victor Meyer Schoenberg and Kenneth Cooker put forward in the book "The Era of Big Data" that "the processing of data in the era of" Big Data "has three characteristics: it is necessary not to sample the whole, but to be efficient. Do not be absolutely precise, be causal. These characteristics are contrary to the concept of news reporting. "5Big data pursues related relationships and ignores the characteristics of causality, making news focus on presenting the interrelationships in news stories, but failing to answer the causal logic relationship behind news stories

In short, the impact of big data on news dissemination is positive and profound. While affirming the value of big data, we must also see that big data is only a quantitative means, and we should not overstate its role. First of all, computers cannot replace people after all. Computers cannot understand the mood, sorrow, and psychological characteristics of the human brain. If we are passively controlled by computer data, we will deviate from the original intention of big data. In addition, the development of big data in the news and media industry must not violate the law of news dissemination, nor can it deviate from the guide of problem solving. Therefore, the value of big data also depends on the use of big data, to follow the law of news dissemination, to solve real-life problems as a guide to improve the level of news dissemination.

Acknowledgement

This paper is the research result of the Philosophy and Social Science Research Project of the Department of Education of Hubei Province in 2019,"Research on the Transmission Path of Wuhan City Image under the background of Mobile Internet" under grant 19G074. Thanks for the support of Education Department of Hubei Province and Wuhan Donghu University.

REFERENCES

[1] Liu jianming, Big data is not a panacea, Beijing Daily, May 6, 2013.

[2] Liu Changle., Data has become the new core competitiveness of companies and countries., Xinhuanet., July 9, 2013.

[3] British "Guardian", Twitter outweighs disadvantages in riots, Sina Technology, August 26, 2011.

[4] Zhu Jianhua:, Data-Driven News-Media's Core Competitiveness in the Big Data Era, People's Network, July 18, 2012.

[5] Liu Yikun, Data Journalism Production in the Big Data Era: Status, Impact and Reflection Modern Communication (Journal of Communication University of China), November 2014.