

Hotspot Analysis of Big Data Research on China's Library and Information

Guanchun Xu

Gannan Medical University
Ganzhou, China 341000

Wenjun Cai

Gannan Medical University
Ganzhou, China 341000

Abstract—The rapid development of computer and information technology has accelerated the progress of China's modernization construction; new concept, new theory, new technique, such as, big data, etc., have emerged as the times required, and injected a steady stream of development vitality into all fields and industries of society. In recent years, the field of library and information has become the focus of social attention, and the hot effects produced by big data are also increasing year by year. Therefore, this paper takes the massive database of "knowledge Network" as the research object, and comprehensively discusses the hot search information of "big data" to explore the research focus and hot issues of big data in the field of library and information in China.

Keywords—library and information field; big data; research hotspot

I. INTRODUCTION

Quoting big data's conceptual content in the Big Data Era written by Viktor Mayer-Schoenberg and Kenneth Cukier, it can be summed up as five features, that is, Volume, Velocity, Variety, Value, Veracity. The application time of big data's concept is relatively short in the field of library and information in China, which is mainly due to the previous data collection, which takes a lot of manpower and time to aggregate. Up to now, the database in the field of library and information has only first seen the scale. In view of big data's research hotspots, authoritative, serious and sophisticated academic institutions should be taken as the object of data source; and the HowNet database has the above characteristic. Therefore, it has also played a positive role in promoting the precision of big data's research hotspots with the help of "HowNet" database.

II. THE MAIN SOURCE OF DATA

This paper mainly takes the database of HowNet as the data source object, takes the three words of "big data" as the key words with the help of the simple search interface of HowNet, inputs the search bar, and then analyzes, studies, arranges and sums up the searched data information. Among them, the literature that does not contain the keywords of "big data" and the literature containing only one or two keywords are not regarded as the research objects. When these non-research objects are eliminated, a total of 1027 academic papers are searched. In view of these documents, this paper classifies and statistics the research hotspots of big

data by using the methods of content analysis and measurement. Of which, the content analysis method is to carry on the accurate analysis to the common word, carries on the effective statistics to the frequency of "big data", and then carries on the thorough research and the analysis to the retrieval data with the help of the statistical analysis tool. With the help of the combination of mathematical knowledge and statistics, the econometric method analyzes the effective utilization rate and circulation rate of the literature to understand and master the external parameters and macro-laws of the literature related to the keywords of big data.

III. QUOTING THE PROBLEMS OF DATA

A. The Value of Data Application Is Not High

With the arrival of big data's era, our life is always bordered with big data, and people's work, study and academic research are closely related to big data. It is the existence of this phenomenon that "huge amount of data" dominates network resources, which makes some data documents lack of authority and lack of research value. By retrieving big data keywords in the HowNet, it is not difficult to find that the data matching the keyword is numerous, but the similarity rate of the data is very high. Some of the data are completely the fake products engraved by a mold. Especially for those who participate in academic research and scientific research, it is impossible to select a paper with reference value and correct guidance from a large number of literatures. Some scholars even store some original data as confidential materials, which cannot be extracted by retrieval method. Therefore, these searched papers also lose their practical reference value.

B. Increased the Frequency of Retrieval and Lack of Original Data

Through the retrieval in the "HowNet" database, it is found that a large number of papers and documents contain only the introduction of vague concepts, general keywords, author signatures, publishing organizations and so on. If the people who need the relevant literature cannot find the content that matches their arguments, arguments and topics when they retrieve these contents, they will need to repeat them many times, which not only increases the retrieval workload, but also is likely to be futile. In addition, most of the contents of the searched papers are reorganized and combined by the author. In particular, some data information

ignores the calculation process, and lists the results directly, which makes the literature content lose its original taste. It is even more difficult to retrieve the original data in the massive data.

C. The Arbitrariness of Document Citation Tagging

When writing a paper, the author often refers to the data literature published by the authority, but it often depends on the marking, and it is clear that the data literature quoted comes from the original data. However, the author confuses one thing with another in the paper, which makes it impossible for researchers to find and correspond to the relevant raw data, and cannot prescribe the right medicine. As a result, a document with solid original data has lost its reference value. According to incomplete statistics, irregular references account for more than 30% of the total number of papers. In addition, some authors of the paper only mark the source of data to the beginning of the article when writing. However, there are no marking marks in the main body, which makes the reference of the paper unable to accurately identify the source of the original data, and there are many doubts about the authenticity of the data.

IV. ANALYSIS AND DISCUSSION ON THE RESEARCH HOTSPOTS OF BIG DATA IN THE FIELD OF LIBRARY AND INFORMATION

A. Big Data and Library

At present, the concept of big data has been deeply rooted in the hearts of the people, and big data's practical application has also radiated to every corner of society. But as a social service platform for transmitting scientific and cultural knowledge, the library should also integrate the internal book resources under the background of big data, form a modern service pattern of collecting books with complete types, optimizing service flow and big data efficient retrieval, so that more people can enjoy the great changes brought about by big data to people's lives. The wave of big data has played a positive role in promoting the long-term development of the library, because big data's massive storage function can meet the needs of different strata for book resources. When the reader walks into the library and applies big data technology to supply knowledge, it accelerates the development process of library informatization and digitalization.

B. Big Data and Service

The emergence of big data's concept is completely based on the rapid development of Internet technology, and big data, as the service resource of Internet transmission terminal, is being displayed in front of the public with the concept of modern information service. In view of the literature and materials of the searched papers, we can see that the personalized and specialized knowledge is relatively lacking, and the knowledge structure is not rigorous, and the knowledge content is mixed. But big data's service value and effect is based on changing the current chaotic resource pattern, and establishes a new information service model which meets the requirements of the development of the times and the needs of public knowledge, so as to completely

change the unscientific, unreasonable and irregular big data original model in the past.

C. Big Data, Library and Information

In the field of library and information, big data's theoretical content must also be put on an important agenda. Only by mastering big data's research concept, research core, development context and future trend, we can combine theory with practice, make better use of big data, and integrate more resources with positive value for the field of library and information. The emergence of big data has added a new data paradigm in the field of library and information. Data retrieval personnel only need to apply this paradigm to obtain more valuable information content in time. In addition, the appearance of big data has also changed the methods, tools and research objects in the field of book intelligence, and also broken the past structure of book intelligence, and then transformed into current unstructured, and the research method of information intelligence is more professional.

D. Big Data and Resources

In the field of books and intelligence, resource construction is the primary task, and resource construction is a basic support work, especially, the normal operation of library and information institutions and the full play of the functions of all kinds of service platforms must rely on the achievements of resource construction. The application of big data is also to embody the value of resource utilization incisively and vividly, and to raise the achievements of resource construction to a new height. At present, there are many types of data resources involved in the library and information institutions in our country, which not only cover astronomy, geography, humanities, science and technology, culture and education and other resources, but also integrate some non-natural disciplines into the data resources. These resources not only provide a better resource base for scientific researchers, but also make the resource service system of the whole society more and more perfect. This is the product of big data's era. Therefore, library and information institutions should be fully aware of the importance of optimal allocation of resources, and integrate all the available resources and valuable resources contained in big data's technology to meet the needs of users and the needs of society.

V. CONCLUSION

Big data's research hotspots also include many contents in the field of China's library and information. However, the field of library and information has realized seamless connection with big data on the current development direction of big data. Although the information and information resources provided in this field are not perfect enough, the authority of the content is poor, the accuracy of citation literature is low, and the query of original data is difficult, this technical difficulty will be solved with the unremitting efforts of research staff of big data. The research hot spot of big data in the field of library and information in our country is bound to become the real hot spot and the focus issue closely related to the people.

REFERENCES

- [1] Si Li, Wang Yuwa. Evolutionary Analysis on big data's research in the field of China's library and information [J]. *New Century Library*, 2018(12):88-94.
- [2] Zou Yi. Hotspot analysis of big data research in China's library and information [J]. *Market Conditions*, 2017(39):163.
- [3] Zhou Quan. The Excavation and Importance of the Hot Field of Library and Information driven by big data [J]. *Youth*, 2019(2):295-296.