Research on the Development of Traditional Literature Database Based on Big Data

Xiaoxiao Liu¹*, Xinhua Liu²

¹Shanxi Institute of Socialism, Taiyuan, 030000, Shanxi, China
²College of Bioresources Chemical and Materials Engineering, Shaanxi University of Science & Technology, WeiYang District, Xi’an, 710021, Shaanxi, China

*Corresponding author (E-mail: shuixiaoer1985@163.com)

Abstract: With the explosive growth of Internet data, the traditional Internet data collection and management mode cannot fully meet the requirements of data management and utilize in the era of big data, and the mining and utilization of big data will completely solve this data management problem. The use of big data has brought great convenience to the research and mining of contemporary Chinese traditional culture, as well as the development of communication. Based on these big data and technologies, we can effectively establish the database of Chinese traditional culture. This paper studies the development of traditional literature database based on big data. In order to better study the development of traditional literature database, this paper puts forward the method of integrating big data technology with the development of traditional literature database. By analyzing the current situation of traditional literature database, a set of traditional literature database suitable for the development of the current era is formulated. Through the analysis, it is found that the method proposed in this paper has important practical significance for the research on the development of traditional literature database based on big data and the inheritance of Chinese traditional culture.

Keywords: Chinese traditional culture, Big Data Technology, Traditional Literature, Database Development
1. Introduction: The significance of Traditional Literature Database Development Based on Big Data

Traditional literature is an important resource to study Chinese traditional culture, it also belongs to the discipline of Chinese traditional culture [1, 2]. Fully mastering historical materials is the basis of studying Chinese traditional culture. As Mr. Fu xuancong said, "ancient books, including ancient books and historical books, are related to Literature and historical materials." Therefore, how to have so many historical materials is a problem. In ancient times, to be able to "read ten thousand volumes" was considered enough for studying of traditional culture, even was considered as the whole information a person has; today, this situation has changed, and it is far from enough for one person to read 10000 volumes of books. Similarly, today's Chinese traditional culture lovers are no longer satisfied with "reading 300 Tang Poems". To meet their needs, it is not enough to talk about 300 Tang poems.

The effective way to solve this problem is to establish the traditional literature database on big data. "Big data" will definitely be listed if we want to choose the popular technical terms in the field of science and technology in recent years. The word "big data" [3-5] is not only popular in the field of science and technology, but also frequently appears in the field of social science and the humanities. This also provides an opportunity for the development of traditional literature database. This new word has too many meanings and contents, which has been hotly discussed and widely quoted by the media. At the same time, different professional groups in different fields have given different meanings according to their own cognition and understanding of the word. However, no matter which field it appears in or from which perspective, the word big data appears frequently, which indicates that the development of information technology has changed from traditional calculation to data collection and statistics. Database [6-8] is closely related to "big data" and plays an important role in data management. This paper discusses and analyzes the research progress and trend of big data interpretation based on database vision.

This paper studies the development of traditional literature [9, 10] database based on big data. The development theme of the current era is the development of science and technology. We are in a big data era, so the development of anything cannot be separated from the help of big data technology. The traditional management method of literature database has been unable to meet the needs of the development of the times. Therefore, this paper puts forward the idea of using big data technology. The method of integration with the development of literature database makes a precise analysis of the current situation of the development of traditional literature database, so as to formulate a new road suitable for the development of traditional literature database in the new era. Through the analysis, it is found that the method proposed in this paper has important practical significance for the research on the development of traditional literature database based on big data.

2. Research Methods of Traditional Literature Database Development Based on Big Data

2.1 Big Data Technology

In the near future, big data technology will enter everyone's life, and gradually become familiar with most people. It will also become a contemporary technical support for the inheritance and creation of
Chinese traditional culture. With the help of big data technology, Chinese traditional culture can realize "old wine can be bottled new", to enhance the soft power of culture with modern science and technology, to inject science and technology into cultural development, greatly expanding the space and mode of cultural and its cultural industries development. Besides, the existence of cloud computing provides a strong support for the processing of big data, and the operation of big data is inseparable from cloud computing. With the deep integration of new generation cloud computing and traditional big data technology and the further arrival of big data era, cloud computing will enable culture and related industries to better continue to closely integrate with traditional cloud computing and big data, and the mutual fusion relationship between them will be more tacit and closer, so that big data can play a better role of mutual integration and linkage in enterprises. Cloud computing and big data will be more and more understood, known and applied by the public, and become the social topic and focus closely paid attention to by various industries in recent years. With the rapid application and development of cloud computing and big data, for all walks of life, science and technology will trigger a greater information revolution and change the patterns of people's lives and cognition.

2.2 The Development Method of Traditional Literature Subject Database

The so-called data refers to the symbolic record of things. As an entity defined as meaning, it has certain internal relations with the existence form of things. Generally speaking, the development of topic database is realized through data mining, processing and analysis in the process of data evolution. Before the 1950s, the content of data management was very simple. Because the amount of data is very small, it is possible to input a large amount of data only when the business needs it. In the 1960s, with the further birth and development of database system, the database management system has entered the design stage of subject database system. With the increasing amount of data in the management system, a large number of data in the system began to show a decentralized state, so we put forward the concept of the so-called subject type database system. The so-called subject type database system is a kind of technical type database which is designed and planned for the purpose of defining business objectives and themes. By reducing and increasing the number of databases and related technology applications, we can effectively solve the problem of subject data redundancy and low sharing rate of related data. In recent years, the trend of data diversification and explosive rapid growth makes the databases of various disciplines close to the advanced big data technology, and the methods of data resource management and business planning of subject databases must be further changed.

3. Experimental Correlation Analysis

3.1 Experimental Background

The accumulation of data is conducive to mining deeper information. For example, in the current national trial credit reporting system, most of the data selected by the first batch of enterprises are provided by Internet companies, which not only shows that Internet start-ups hope that more users can see the relevant data. However, different Internet start-ups can conduct comprehensive analysis on relevant technologies and data based on their own data, and accurately score users' behavior, relevant technology and credit. Similarly, government departments dare to use the Internet start-up company data related to these data to express doubt and approval of the reliability of the relevant technology and
3.2 Experimental Design

Data analysis can also be used for annual leave and Spring Festival passenger flow to prevent possible problems in advance. There is also the analysis based on user retrieval data, such analysis can see the attention of online people, as well as the changes in the popularity of related issues, can pay attention to the change of public opinion, and solve related problems in a timely manner. The use of big data is becoming more and more popular, and the value of big data is also constantly reflected. This paper investigates the income structure of China's major network companies in 2019, the survey results are shown in Table 1:

**Table 1.** Advertising revenue and total revenue of domestic network companies in 2019 (unit: 100 million yuan)

|        | Advertising revenue | Total revenue | percentage |
|--------|---------------------|---------------|------------|
| Baidu  | 2663                | 2980          | 89.3%      |
| Sina   | 1432                | 1765          | 81.1%      |
| tencent| 1222                | 18956         | 6.4%       |

4. Discussion

4.1 Analysis of the Research Status of the Development of Traditional Literature Database Based on Big Data

With the development of modern information internet technology, the application and popularization of mobile terminal network communication equipment, social network, e-commerce and other websites easily have more than one million server-side customers. Whether the mobile terminal can effectively solve the bottleneck and pressure of the rapid growth of mobile network server-side website users has a very decisive impact on the sustainable survival and development of the website. Therefore, when a large number of customers visit their website in mobile terminal, the website needs to handle the concurrent requests of customers. How to reduce the waiting time of customers and improve the quality of the website is particularly important. In order to shorten the waiting time of user requests, it is necessary to dynamically enhance the ability of database to process concurrent requests. At present, the domestic large-scale website is through the establishment of server cluster to alleviate the problem of rapid growth of users and visits. However, the data throughput per unit time is still limited due to the unchanged read-write performance of traditional databases. Based on the survey and analysis of the major instant messaging software industry data in China in 2019, the survey results are shown in Figure 1:
As shown in Figure 1, in general, WeChat, QQ and Alibaba Wangwang have the largest average daily coverage, with 523.24 million, 221.34 million and 93.56 million respectively. The number coverage of FeiXin in 2019 is far less than the above three instant messaging software, only 2965.9 million people. Because the population base of our country is relatively large, the daily average number of people covered by this instant messaging software is counted by 100 million units, which poses a great challenge to the database carrying capacity of China's network companies.

Take a typical example. The classification of tourist attractions (tourism destination level) is restricted by tourism resources and market attractiveness. The attraction of tourism resources refers to the resources with ornamental and entertainment value, historical value and cultural value. Market attraction refers to the popularity, reputation and market radiation ability of scenic spots. Tourism literature works belong to the category of literature tourism resources. The number and spatial distribution of poems in different tourist destinations are visualized by using hierarchical statistical graph method and the relationship between tourism literature and tourism destination is analyzed. The number of Li Bai's and Du Fu's poems and tourist destination data in China are statistically analyzed; the statistical results are shown in Figure 2:
Figure 2. Statistical analysis of the number of Li Du's poems in different tourist destinations

The higher the number of tourist destinations, the lower the number of tourist destinations. This roughly shows that the development of traditional literature library cannot be separated from the foil of traditional literature, through the vigorous promotion and publicity of traditional literature, the traditional literature library can be effectively developed.

4.2 Suggestions on the Development of Traditional Literature Database Based on Big Data

Big data has brought great convenience to the research and dissemination of Chinese traditional culture. With the new resources created by the database, people can break through the limitation of time and space and find more information in a short time. At the same time, the database also provides researchers with new methods and new ideas, and provides new channels and platforms for Chinese traditional culture lovers. Using the database, researchers can open up new research fields and carry out more scientific research. Traditional culture lovers can also have a more comprehensive understanding of traditional culture and carry out the secondary dissemination of traditional culture.

The concept of topic database has been proposed for some time, but the related theories and methods of topic database have not been applied to practice on a large scale, so there are still many problems worthy of further discussion. With the advent of the era of big data, new theories and problems will inevitably arise in thematic databases. The data under the background of big data is mostly unstructured data, and the data quality is uneven. In the background of this era, the topic partition method in topic database needs to be improved. How to establish a reasonable model or calculate the relationship between entities and entities, between entities and activities, quickly divide the massive data into reasonable data units, and establish the subject database that meets the business
requirements is the primary problem in front of us.

At present, big data is still in the preliminary exploration stage, and there are not many research results on big data management and decision-making. The establishment of subject database is for data management and decision-making through data. Therefore, subject database under the background of big data should focus on data analysis and data management decision-making of discipline database. According to the characteristics of some big data, which features have the greatest impact on data management decision-making, and how to communicate and cooperate between users in different fields and researchers in different disciplines can help solve the problems related to the subject database under the background of big data, which will be the next research direction.

From the development trend of database, the decline of traditional database is irresistible. In the face of the current amount of data and the analysis and application of data, the traditional database has been unable to meet the needs of all kinds of industry include cultural industries. As a result, people turned to other solutions. These database solutions can complete the application that the traditional database cannot complete, and can also achieve the performance indicators that the traditional database cannot achieve, but there are still some defects that industries cannot accept. Therefore, according to these solutions, people constantly improve the database solutions, so that the database integration machine appears in front of people. This provides more advanced technical support for the establishment of traditional literature database.

5. Conclusions

This paper mainly introduces the development and research methods of traditional literature database based on big data. With the continuous development of society, human's scientific and cultural achievements have reached a high level. A large number of scientific and technological forces have helped people improve the efficiency of work and learning. Now the development of all walks of life cannot do without the analysis of big data technology. In order to make the development of traditional literature database more smoothly, this paper puts forward the method of integrating big data technology with the development of traditional literature database, analyzes the development status of traditional literature database accurately, and formulates a set of development path more suitable for traditional data base in the new era. Through the analysis, it is found that the method proposed in this paper has important practical significance for the research on the development of traditional literature database based on big data.

References

[1] Zhang Y, Min Q, Zhang C, et al. Traditional culture as an important power for maintaining agricultural landscapes in cultural heritage sites: A case study of the Hani terraces[J]. Journal of Cultural Heritage, 2016, 25:170-179.

[2] Zhang P. Good gibbons and evil macaques: a historical review on cognitive features of non-human primates in Chinese traditional culture[J]. Primates, 2015, 56(3):215-225.

[3] Assuncao M D, Calheiros R N, Bianchi S, et al. Big Data computing and clouds: Trends and
future directions[J]. Journal of Parallel & Distributed Computing, 2015, 79-80:3-15.

[4] Hashem I A T, Yaqoob I, Anuar N B, et al. The rise of "big data" on cloud computing: Review and open research issues[J]. Information Systems, 2015, 47(jan.):98-115.

[5] Lv Y, Duan Y, Kang W, et al. Traffic Flow Prediction With Big Data: A Deep Learning Approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2):865-873.

[6] Groom C R, Bruno I J, Lightfoot M P, et al. The Cambridge Structural Database[J]. Acta Crystallogr B Struct, Cryst Eng Mater, 2016, 72(2):171-179.

[7] Ward F, Kerrie V, Nima S, et al. The Immune Epitope Database: How Data Are Entered and Retrieved[J]. Journal of Immunology Research, 2017, (2017-5-29), 2017, 2017:1-13.

[8] Cuticchia A J. Future vision of the GDB human genome database.[J]. Human Mutation, 2015, 15(1):62-67.

[9] Gurney R M, Caniglia B S, Mix T L, et al. Native American Food Security and Traditional Foods: A Review of the Literature[J]. Sociology Compass, 2015, 9(8):681-693.

[10] Latulippe N. Situating the work: A typology of traditional knowledge literature[J]. Alternative An International Journal of Indigenous Peoples, 2015, 11(2):118-131.