Chinese National Characteristic Historical Database System Platform Based on Big Data Technology

Quan Qiu(✉)

Science and Technology College, Gannan Normal University, Ganzhou 341000, Jiangxi, China
54099963@qq.com

Abstract. In order to carry out in-depth research on historical data with Chinese national characteristics more conveniently, it is necessary to use big data technology to collect, organize and analyze historical data resources. Big data technology can digitize various forms of data, describe and manage data in a unified manner, and achieve efficient use of data. The historical database constructed in this paper uses the migration process model to provide users with a clear human development route according to the temporal and spatial characteristics of people's migration in history. This database system platform can provide learners and researchers with a reliable way to query historical data, which is very practical for the protection and inheritance of history.

Keywords: History · big data technology · national characteristics · database

1 Introduction

As a country with a history of more than 5,000 years, China has produced many intangible cultural heritages in the course of its development. China's intangible cultural heritage records the economic, cultural, political, religious and other historical and cultural information formed in the course of Chinese and human development. In order to preserve this information and use it for inheritance, China has used many methods to protect it, such as libraries, archives, cultural centers, historical museums, and so on. Nowadays, with the development of computer technology, computer technology occupies an important position in people's life and work. In order to improve the dissemination efficiency of Chinese national characteristic culture and facilitate people in different regions to understand Chinese characteristic culture, this paper constructs a Chinese national characteristic historical database system based on big data technology. In this system, the original historical data in paper, video, audio and other modes are collected, organized, analyzed and shared, and provided to users with intelligent resource services, bringing greater economic and social benefits to the society.

Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 1099–1105, 2023.
https://doi.org/10.2991/978-94-6463-034-3_112

2 The Significance of Developing a Historical Database

2.1 Conducive to Academic Research

Due to the different levels of development in different regions of China, the historical materials of some ethnic groups are recorded through oral dictation, and most of these historical materials are provided to researchers for analysis. The application of big data technology in historical mining can summarize the history of national development for researchers and explore the laws behind historical development. Big data technology can collect a large amount of data, and the analyzed data can not only provide valuable information for history, but also provide research materials for various disciplines such as ethnology, linguistics, library and information science, and political science. Big data technology can perform cluster analysis on a large amount of data, and through the identification of audio, the collection of text data, and the mining of popular historical knowledge, it can point out the direction for academic research.

2.2 Protecting Cultural Traditions

Big data technology can analyze the data structure of historical data and user behavior, and the results of data analysis can help protect and inherit Chinese traditional culture. Big data analysis can analyze historical and cultural materials, find out where historical and cultural gatherings are located, select key protection contents and protection objects, and then take corresponding protection measures for these regional cultures that need key protection according to the results of data analysis. Nowadays, the Internet is very popular, and big data technology can be used to collect users' access points, search keyword frequency, hot resources, etc. in the network. Big data technology can understand the proportion of different types of historical and cultural resources based on resource statistics, control each type of resource data, and collect data. From the theme, region, quantity, etc. of historical and cultural resources, the distribution law of resources can be found. After understanding the distribution law, you can collect data accordingly, and optimize and update resources regularly.

2.3 Develop Big Data Resources

The continuous development of big data resources by the historical and cultural system can further promote the development of big data technology and improve the utilization of big data. The system can obtain the access traffic according to user access data statistics. Based on these data, the database can be improved and the retrieval strategy can be optimized. The system can improve the retrieval database and the retrieval database through continuous machine learning methods, and automatically push popular resources. Big data technology can also discover associations between data and establish semantic retrieval strategies. Big data technology can innovate service methods, collect and sort out the original scattered resources, and improve the utilization rate of resources.

Table 1. Data formats stored in this database

Resource Type	quantity	Proportion (%)	total
audio	985	22.9	4292
video	568	13.2	
Historical data (text)	2523	58.8	
picture	125	2.9	
Others (relevant books, journal articles, etc.)	91	2.1	

3 Data Sources and Analysis Tools

The data source of this article is the national cultural data information on the Internet, including audio, video, text, pictures, etc. The data sources in the database are mainly digital libraries in various places, and the data resources in these libraries have high credibility. The content of the data includes historical data on languages, ballads, literature, education, philosophy, etiquette, etc. of all ethnic groups in China. The database arranges and organizes these data in various ways, including dates, languages, regions, dialect types, themes, keywords, sources, formats, etc. After entering the database, users can choose the format of the query data according to their own needs (Table 1).

The database also analyzes the data. The data analysis of this database is mainly to summarize and mine the source and format of the data. At the data analysis level, analysis tools are generally divided into four layers, namely data storage layer, data reporting layer, data analysis layer, and data presentation layer. Each layer is divided into portal level, department level, enterprise level and BI level. In this database, the storage layer uses the MySQL database. MySQL database is one of the most common relational database relational systems today. The combination of this database with PHP and Apache can provide a good development environment for the database. In order to show users more intuitive data analysis results, the system uses visual analysis tools. This database uses the department-level software Tableau, which has the function of visual analysis, which can visualize the results of data analysis and display the results of data analysis more intuitively.

4 Database Structure

4.1 Basic Structure

The spatial data and attribute data in the database are the key to the data model. There are two main relationships between the spatial data and the attribute model, namely merged storage and associative storage. Chinese national cultural resources are rich and varied, and there are many forms of expression, so the data model used in this database adopts the method of setting up associations between separate storages. In the database, spatial data is stored independently to avoid data redundancy caused by overlapping spatial data. The historical and cultural attribute data are stored according to categories, which

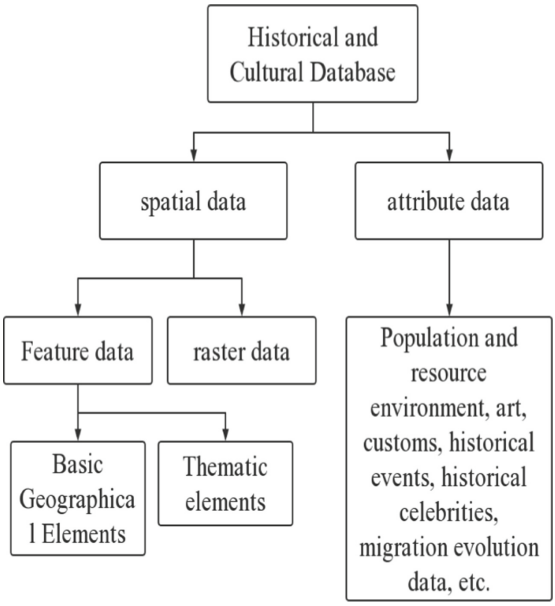

Fig. 1. The structure of the historical and cultural database

is beneficial to the storage and management of the data by the database, and it is also more convenient for users to conduct special research (Fig. 1).

The data model is the main support and architecture of the database system. Historical and cultural materials are closely related to spatial location. In historical and cultural data, geospatial location is an important description of information data. Therefore, in this database, the spatial geographic information of the information data will be described and recorded in detail.

In the course of human development, the geographical location of human beings will continue to change over time. When designing the historical and cultural data model, it is also necessary to consider the combination of time-varying features and spatial-varying features. In the process of storing, expressing and managing data information in the database, it is necessary to focus on the spatial expression and temporal and spatial correlation of information data. Figure 2 shows the spatiotemporal data model in this system.

4.2 Model of Human Migration Process

In this system, a model is set for the migration process of ethnic groups, as follows:

$$\text{Process}(i) = f(\text{ST}(i), \text{ET}(i), \text{L}(i), \text{Info}(i), \text{R}(i)) \ (i \in [1, 5])$$

In the formula, Process(i) refers to the i-th migration process. ST(i) is the start time of the ith migration process. ET(i) is the end time of the i-th migration process. Info(i) is migration process information. L(i) is the ith migration gradient. R(i) is the migration process remark.

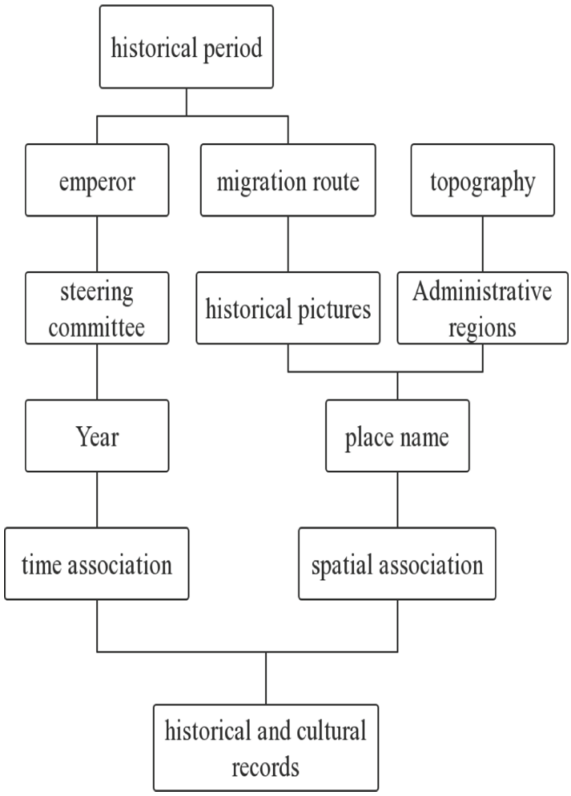

Fig. 2. Historical and cultural spatiotemporal data model

Through the migration model, the regions involved in the migration can be associated according to the order of time.

5 Conclusion

The historical and cultural database can provide basic data support for the study of Chinese history, and can protect and promote traditional culture. This paper uses big data technology to collect and sort out some ethnic historical and cultural materials, and build a complete historical database. This database can be developed into a knowledge-based website for the public in the future, providing more people with access to historical data.

Bibliography

1. Bi Jiaqi, Li Chunmei. Talking about the process evaluation method of historical research study - portfolio evaluation method [J]. China Youth, 2017(02): 276-277.
2. Chen Shaogen. Research on the method of integrating archives and database construction based on the surveying and mapping archives management platform [J]. Surveying and Mapping and Spatial Geographic Information, 2021, 44(11): 114-117.
3. Fu Shihan. Archival value mining and archival propaganda from the perspective of historical research [J]. Lantai Neiwai, 2019(29):63-64.
4. Gao Weibo, Pang Wenjing, Chen Beibei, Li Zhongqin. Information Security Design and Implementation of Geological Data Management System during Construction [J]. China Mining Industry, 2021, 30(S2): 136-142.
5. Guo Haixia, Zeng Gang. Knowledge Graph Analysis of Hot Spots and Frontiers of Ethnic Cultural Identity Research: Based on CSSCI (1998-2020) Data [J]. Journal of Southwest University for Nationalities (Humanities and Social Sciences Edition), 2021, 42(04) :223-233.
6. Jiang Zhujun. Suggestions on the restoration of collection materials and database construction in the library and database of the geological and mining industry—taking the Geological Archive of Gansu Provincial Bureau of Geology and Minerals as an example [J]. Gansu Geology, 2021,30(04):90- 93.
7. Mu Jinrong, Zeng Yuqing, Jin Xiaoyan, Liu Xiaoe. Exploration and practice of urban geological data information service cluster technology [J]. Land and Resources Guide, 2021,18(02):86-91.
8. Tian Zhiyong. Establishment and application of traditional village survey data database [J]. Rural. Agriculture. Farmers (version B), 2020(08):37-38.
9. Wang Yanchun. From the perspective of historical research, the definition of the filing scope of documents and materials in grassroots units and the definition of archives storage period [J]. Archives World, 2018(05):41–42+40.
10. Xi Nan. Design and implementation of the database and application system of modern Chinese science and technology associations based on knowledge maps [D]. Beijing University of Posts and Telecommunications, 2021. DOI: https://doi.org/10.26969/d.cnki.gbydu.2021.001912.
11. Yu Guoqian, Tao Guangyi, Zhao Tianyu. Research on key technologies of unstructured hydrological data database structure standards [C]//. Proceedings of the 2021 (Ninth) China Water Conservancy Information Technology Forum., 2021:632–636.DOI :https://doi.org/10.26914/c.cnkihy.2021.006884.
12. Yuan Bolan, Huang Li. Transformation and basic characteristics of contemporary ethnology and ethnic culture research paradigm [J]. Ethnic Forum, 2021(04):41-48.DOI:https://doi.org/10.19683/j.cnki.mzlt.2021.04.002 .
13. Zhan Yike, Shi Linyu. “Panoramic Records”: The Reflection and Practical Impact of Archival Thinking in the Transformation of Social and Historical Research [J]. Shanxi Archives, 2022(01):95-102.
14. Zhao Lan. Construction of Active Fault Detection Database in Sanhe City [D]. Institute of Disaster Prevention and Technology, 2021. DOI: https://doi.org/10.27899/d.cnki.gfzkj.2021.000034.
15. Zhu Aiqing. The value realization of oral archives from the perspective of historical research [J]. Office Business, 2020(05):118+120.