Research on the Application of Computer Big Data Retrieval Technology and Cloud Computing in Internet Course

Peisheng Li

Shandong Vocational College of Science & Technology, Weifang 261053, Shandong Province, China, lips@sdvcst.edu.cn, Corresponding author

ABSTRACT

The current Internet teaching resource retrieval system has problems such as large information retrieval errors and low work efficiency. In order to obtain ideal Internet teaching resource information retrieval results, the article designs an Internet teaching resource information retrieval system based on big data analysis technology. The article collects a large amount of data from digital Internet teaching resource information retrieval, and introduces big data analysis technology to establish an Internet teaching resource information retrieval model. Finally, specific examples of Internet teaching resource information retrieval are used to analyze the superiority of the proposed model. Experimental results show that this technology is suitable for organizing, managing and querying massive amounts of Internet teaching resource data, and the query performance is better than traditional relational databases.

CCS CONCEPTS

• **Computing methodologies**; • **Modeling and simulation**; • **Model development and analysis**;

KEYWORDS

Computer big data, data retrieval, internet teaching, data processing, simulation experiment, retrieval system

1 INTRODUCTION

The gradual maturity of the application of computer technology in the field of education has brought earth-shaking changes to the traditional teaching model. In today's digital information age, traditional education models can no longer meet people's requirements for learning methods. The network teaching model is an inevitable product of the information age, and it is also the focus and direction of the future development of the education field [1]. At present, the online teaching model has been widely used all over the world,

especially in some Western countries such as Britain and the United States, taking the lead in reforming and promoting the teaching model. However, due to the late start of the development of some infrastructure and key technologies, and the lack of attention by the national government to them, the development of China is relatively slow, and there is still a certain distance from the advanced and frontier countries in the world.

As a new type of teaching model that is the key development in the education field in the information age, distance education optimizes the traditional teaching model, integrates a wealth of teaching resources, and improves teaching quality and teaching efficiency. Among them, the network teaching platform is the core of the entire distance education system, and the rationality and operability of its design have a vital impact on the structure and function of the entire system. At present, in the international market, many companies have developed various types of online teaching platforms, such as learning Space, Top class, Black board and other commercial platforms and Moodle, Autor, Claro line, etc. as the main representative's Free platform system, and all have good compatibility, can support multi-language systems, so it has a wide range of applications in the world.

The Chinese language teaching system is mainly for people who learn Chinese language and culture. It can centrally manage Chinese language teaching resources and is a platform for Chinese language and literature exchanges and learning. The system is based on VR technology and uses 3dsMax and Unity3D as the development environment. The content that needs to be told in the real classroom can be displayed on the user's mobile device truly and quickly, and it can be remotely realized through auxiliary means such as network communication technology and 3D sound effects. Teach. Users can use the VR Chinese language teaching system to complete teaching tasks in a three-dimensional virtual space, and can also construct a three-dimensional virtual teaching scene suitable for teaching activities; the system supports switching between different teaching scenes and switching between various course contents; in addition, The user can select the course they want to choose, the details of learning, and the control of the sound through the gaze function [2]. At the same time, the system sets up the teacher's 3D model in the virtual 3D scene and is accompanied by animation. The teacher will respond accordingly to the teaching content in the classroom and enhance the sense of reality. The system can also call the camera of the mobile device to integrate the real scene with the virtual scene, which is convenient for users to consolidate the knowledge learned in the classroom.

The choice of this article is based on the virtual display technology platform. As an open-source platform, this platform has high compatibility and powerful functions. It is often used in the learning mode of online courses, which fully meets the needs of this article
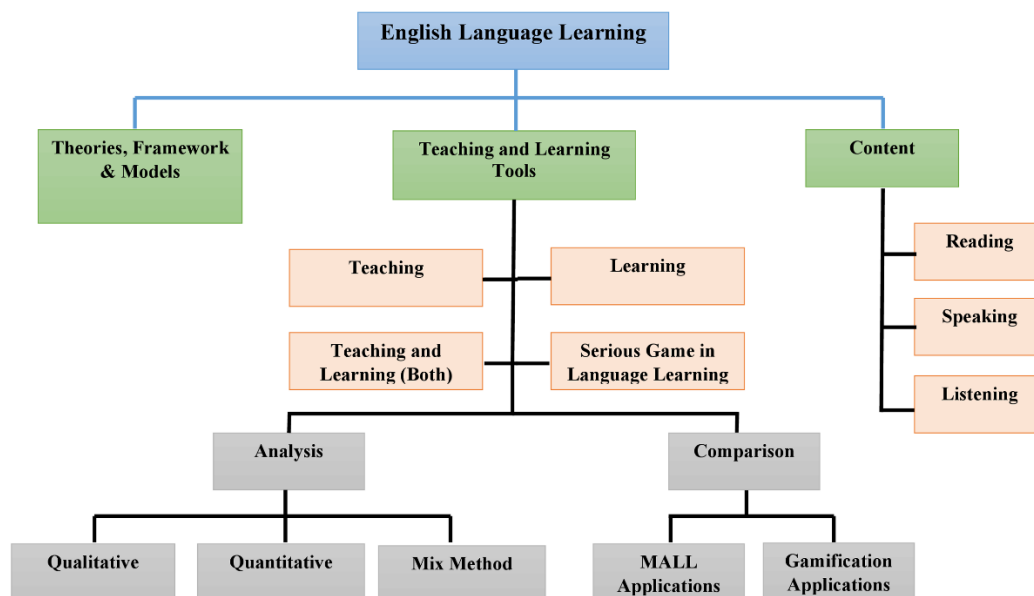
Figure 1: Functional modules of VR Chinese language teaching system.

for network teaching methods. Taking Chinese language and literature teaching as the object, according to the teaching method and functional requirements of the actual teaching mode, the network platform is designed and planned, and the computer virtual display technology network teaching platform is used to greatly improve the learning efficiency of users and enrich the learning methods of students.

2 VR CHINESE LANGUAGE TEACHING SYSTEM DESIGN

The focus of the VR Chinese language teaching system is to use VR technology to create a virtual classroom with a sense of reality and immersion, and to provide users in different geographical locations with high-quality Chinese language education resources. The system uses Unity3D as the development engine, C# language and Play maker plug-in for development, and 3dsMax software to build the model in the scene. The VR Chinese language teaching system is mainly composed of a three-dimensional model module and a human-computer interaction module. The three-dimensional model module is mainly used for modeling and designing the three-dimensional model needed in the system [3]. Through the predetermined modeling plan and drawings, use the extrusion, Boolean, lofting and other technologies in the 3dsMax software to perform three-dimensional modeling of the objects required for teaching, and use the Vary technology to render the relevant models. 3dsMax software has a very powerful modeling function, which can quickly build various models needed in teaching, but its disadvantage is that the model data is often large, which is not conducive to scene integration and network transmission. After the model is built, the established 3D virtual model is optimized through modeling optimization technology to make it more suitable for use in the VR teaching system. The human-computer interaction module is

mainly to design the interaction between the user and the system, based on the Unity engine, and the programming language is C#. The functional modules of the VR Chinese language teaching system are shown in Figure 1

3 SPECIFIC DESIGN OF INFORMATION RETRIEVAL SYSTEM FOR TEACHING CHINESE AS A FOREIGN LANGUAGE

3.1 The hardware design of the system

The learning behavior of traditional teaching information retrieval is multi-factorial and cumbersome, and the purpose of retrieval information is relatively vague. A large number of uncertain factors will lead to semi-structured problems. Based on statistical analysis, traditional teaching multimedia information cannot create a retrieval model, and based on cybernetics, it cannot accurately track teaching resource information. Therefore, the hardware design of the information retrieval system for foreign Chinese teaching resources based on big data is very necessary.

(1) Data source traceability component design

Based on the big data analysis of the foreign Chinese teaching resource information retrieval hardware system, the data source tracking design is mainly to track the multimedia situation of the foreign Chinese teaching resources, and collect user behavior characteristic data in time [4]. This data source comes from the multimedia server terminal memo system, which is automatically saved every 3 minutes, so the data collection error is small, and the data source tracking effect is good.

(2) Data mining component design

The focus of the design of the information retrieval system is to select the knowledge base. The knowledge base is a regular aggregate, and the data mining algorithm based on data entropy can be used to mine different effective data. Assume that the multimedia
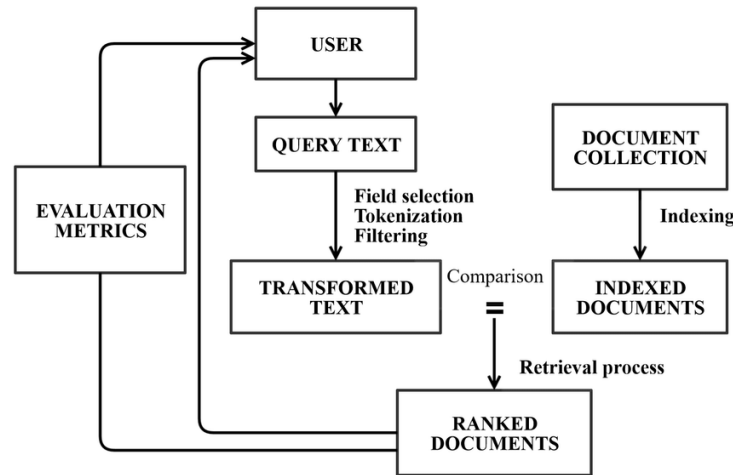
Figure 2: The retrieval process of the retrieval system.

information data system N is:

$$N = \{a_1, a_2, ..., a_m\} \tag{1}$$

In the formula, a refers to the data mining object, if:

$$H_a(N) = -\Sigma_{i=1}^{m} p(a_i) \log p(a) \tag{2}$$

Then, the object p is equivalent to the data mining information of N:

$$R = H_a(N) / H(N) \tag{3}$$

Where: H(N) refers to the information entropy of N; Ha(N) refers to the information entropy of N after the data mining object a. In the process of continuous deepening of data mining, the knowledge base will continue to add new rules to restrict the intelligence of the system. The design of multimedia information mining based on data entropy can not only use the knowledge base to display, but also analyze the mining data results. The retrieval system design is realized by the above three components, and the retrieval process of the retrieval system is obtained, as shown in Figure 2

It can be seen from Figure 2 that users use foreign Chinese teaching resources to retrieve teaching resources. The information retrieval system memo records the user's viewing of teaching resources and collects data. After the data is preprocessed, the data tracking component is transmitted, and the data component uses the data. The mining calculation is fully displayed through the knowledge base, and the data results after the mining are analyzed, and the data is finally transmitted in the system alarm device. If the user is behind the progress of viewing the teaching resources, the system will be warned to start, and the hardware design of the retrieval system will be completed.

3.2 System component development and functions

When developing the system retrieval function, the Hadoop platform is used as the development environment, and three management tools are mainly used: HDFS tools, storage system management and operation related data information; MapReduce tools, all access, downloads, queries, etc. involved in the operation of the system The data is statistically processed, and statistical results are generated, so that the teaching resource administrator can grasp the current situation of the current teaching resource literature access, query and other aspects; the Hive tool uses the information keywords as the management basis to analyze the information resources and store them in the designated folder.

(1) HDFS component functions

The application of this component in system development divides resources into multiple data nodes according to different resource types, and uses control nodes to manage them, so that information retrieval can be implemented in an orderly and efficient manner [5]. Among them, the control node refers to the system administrator, in addition to centralized management of books and documents, according to the needs of document management, a storage space for teaching resource files is organized for users to retrieve and download. For the management of new teaching resources and literature resources, according to the different types of resources, select the corresponding storage path to complete the unified management of file information nodes. The basic structure of the management of teaching resources and literature resources is composed of multiple data nodes, and the teaching resources and literature resources are stored in the form of blocks. By setting the management cycle, resource information is sent to the control node at regular intervals. In actual applications, the user initiates a resource retrieval operation request on the client, and the HDFS component will enable the resource block information transmission function to provide the user with related resources.

(2) Hive component functions

The Hive component is mainly used to analyze and query the resource information of teaching resources. By analyzing the semantics of foreign Chinese teaching resources, metadata can be extracted from it to form an analysis and query operation project execution plan. Run according to this plan, establish a communication connection between the job node and the task node, and complete the mapping task through the execution of the engine program, so as to realize the query of teaching resources and literature resources.
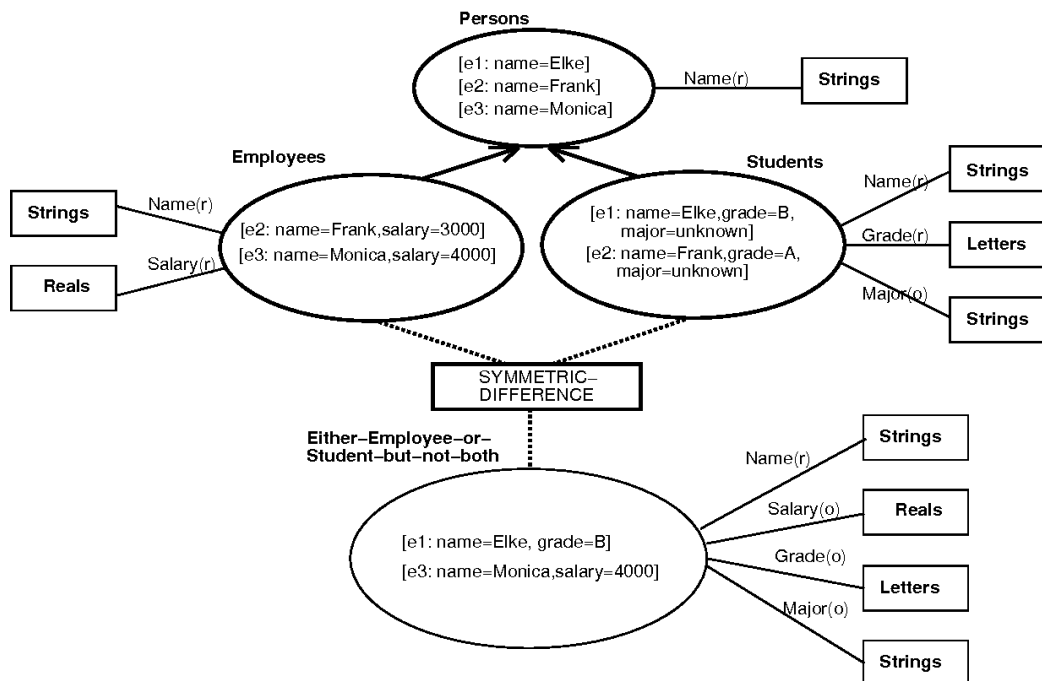
Figure 3: E-R diagram of Chinese language teaching system.

(3) Map Reduce component function.

The MapReduce component is used to develop system information retrieval functions, and use data mining technology to write information mining algorithms to realize cloud retrieval functions. The main operation of this component is scheduling tasks. According to different functions, the system function modules are divided into multiple slices, and each node is responsible for the operation, forming a more complex data mapping relationship, so as to achieve large-area data information retrieval and obtain a more comprehensive search result. The functional development of this component will be introduced in the next part of applied algorithm research.

4 DATABASE DESIGN

The database is the key to storing data information and data processing by users in the software system. According to the user's needs and application functions of the system, a reasonable database data information is built through different data models [6]. For example, the conceptual data model is based on the theory of entity correspondence between objects, that is, the E-R theory, and it is perfected and optimized, and its data information is optimized according to the different functional needs of users; the logical data model is It is the establishment of specific models of the above models to form the model conversion of the logical relationship between the databases.

4.1 Logical Data Model

Convert the conceptual data model into a data model supported by specific database products, such as a relational model, to form a database logic model. You can use the Sybase Power Designer tool

to directly build the logical data model LDM, or convert it through CDM.

4.2 Physical Data Model

On the basis of the logical data model, according to the characteristics and processing needs of the DBMS, physical storage arrangements are made, indexes are designed, and a database model is formed. You can use the Sybase Power-Designer tool to directly build a physical data model (PDM), or you can get it through CDM/LDM conversion. The process of database design is generally divided into several links: data analysis, theoretical design, logic design, implementation design and implementation and testing verification [7]. Different databases use different naming principles. Moodle's database model includes information such as system characteristics, module names, and course classifications in the name of the database. In this way, the basic information of the database can be expressed clearly, and it can also prevent the occurrence of name duplication. When the module information of the system changes, we usually export the Moodle database to an XML file, and then process and analyze each module separately. Figure 3 shows the E-R diagram of the Chinese language teaching system.

5 SYSTEM SIMULATION EXPERIMENT DATA ANALYSIS

For the digital foreign Chinese teaching resource information retrieval model based on big data analysis technology designed in this article, in order to effectively verify its feasibility, 500 texts related to the SaaS field are selected as the foreground corpus, including the cultural field, the news field, etc. This time, the established

Table 1: Test results of traditional retrieval models.

Serial number	Search term	Average weight	Average recall	Accuracy
1	Java	0.72	0.76	0.70
2	IOS	0.70	0.76	0.69
3	C++	0.75	0.77	0.75
4	EPR	0.76	0.75	0.72
5	WPS	0.80	0.72	0.74

Table 2: Test results of the newly created model.

Serial number	Search term	Average weight	Average recall	Accuracy
1	Java	0.85	0.88	0.89
2	IOS	0.83	0.94	0.85
3	C++	0.82	0.96	0.87
4	EPR	0.87	0.93	0.94
5	WPS	0.83	0.97	0.86

model and the traditional retrieval model were selected to test, and whether the retrieved matching results obtained can obtain a very good average weight, recall rate and accuracy rate. To ensure that this experiment is sufficiently descriptive, the detection results of the traditional retrieval model are shown in Table 1, and the detection results of the new retrieval model are shown in Table 2. Through the in-depth analysis of Table 1 and Table 2, it is found that compared with the traditional retrieval model, by using the newly-built retrieval model for retrieval, it is possible to obtain better average weight, recall rate and accuracy rate. The reasons are as follows: First, it is related to the search perspective. When searching through the new model, it is necessary to semantically process all search terms and calculate the relevant indexes of all digital foreign Chinese teaching resource information; second; It is related to querying the SaaS field, and it can obtain high-accuracy extraction and retrieval results.

6 CONCLUSION

The network teaching system based on the Moodle platform as a representative has important guiding value for promoting the application of network technology in the field of education. The design and implementation of the Chinese language distance learning system under the online teaching Moodle platform Fujia model

not only integrates excellent teaching resources, but also enriches teaching methods for efficient education, promotes the educational development of schools, and improves teaching quality and efficiency.

REFERENCES

[1] Yalun, A. International Promotion of Chinese Language in the New Era. International Education Studies, vol.12, pp.67-79, July 2019.
[2] Jiang, X. Analysis of Teaching Optimization of Chinese Language and Literature under the Background of Times of Internet Plus era. Contemporary Education and Teaching Research, vol.2, pp.28-30, January 2021.
[3] Huang, F. Teo, T. & Zhou, M. Chinese students' intentions to use the Internet-based technology for learning. Educational Technology Research and Development, vol.68, pp.575-591, January 2020.
[4] Zhao, L. X. Blankinship, B. Duan, Z. Huang, H. Sun, J. & Bak, T. H. Comparing face-to-face and online teaching of written and spoken Chinese to adult learners: An Edinburgh-Sheffield case study. International Journal of Chinese Language Teaching, vol.1, pp.83-98, January 2020.
[5] Yang, Z. & Feng, B. Design of key data integration system for interactive English teaching based on internet of things. International Journal of Continuing Engineering Education and Life Long Learning, vol.31, pp.53-68, January 2021.
[6] Wang, D. & East, M. Constructing an emergency Chinese curriculum during the pandemic: A New Zealand experience. International Journal of Chinese Language Teaching, vol.1, pp.1-19, January 2020.
[7] Ma, C. & Wang, Y. A Study on the Admission System of Foreign Students in Chinese Public Schools—Focus on Chinese Language Education. Open Journal of Social Sciences, vol.8, pp.139-149, August 2020.