Construction of Three-dimensional Network Teaching Resource Platform in Colleges and Universities under Big Data Technology

Yu Chen[a] , Jing Li [b]

Weifang Engineering Vocational College, Qingzhou, Shandong, 262500

[a]adachen0909@163.com , [b]992276264@qq.com

Abstract. At present, there is a lack of unified planning for the construction and application of online teaching resources in colleges and universities, so that online teaching resources are relatively independent and scattered, and there are problems such as difficult transmission and sharing and low application efficiency. In this regard, this paper starts with the analysis of the construction and application status of online teaching resources in colleges and universities, and puts forward a set of construction scheme of three-dimensional online teaching resources platform to promote the innovative practice of teaching mode in colleges and universities. The platform takes big data technology as the core, uses Hadoop framework to complete the distributed storage and management of online teaching resources, and cooperates with Lucene algorithm tools to complete the design and development of search engines, which are encapsulated into standard Web applications under J2EE technology. Practice shows that the platform effectively solves a series of problems in the integration, circulation, application and management of online teaching resources in colleges and universities, expands the application dimension and improves the application efficiency, thus making a beneficial attempt for the construction of smart campus in colleges and universities under the big data environment.

Keywords: big data technology; three-dimensional teaching resources; Hadoop framework; Lucene algorithm; software application

1 Introduction

With the in-depth implementation of "Education Informatization 2.0 Action Plan", digital technology has become the fundamental force to build a modern education system, and it is also an important engine to promote the high-quality development of higher education in the new era. [1] Digital technology can empower the whole area of higher education, and trigger the multi-dimensional transformation and upgrading of school-running thinking, educational concepts, teaching resources, teaching contents, teaching forms and teaching tools. Among them, online teaching form and online teaching resources are the most popular, and they are closely related and complement each other, which points out the direction for the transformation and upgrad-

ing of college education and teaching mode. However, in the actual application process, the construction of online teaching resources in colleges and universities often lags behind the online teaching application. On the one hand, it is because of the long production cycle and high input cost of online teaching resources. On the other hand, the construction methods and content standards of online teaching resources are not uniform in different professional disciplines, and there are certain academic barriers, which leads to problems such as low utilization rate of online teaching resources and poor sharing and transmission. [2] In view of this, this paper believes that under the background of digital intelligence era, colleges and universities should integrate new teaching concepts and teaching technologies, design and develop a three-dimensional network teaching resource platform adapted to the modern education system, and then create a new ecology of higher education and teaching. [3] The three-dimensional network teaching resource platform takes Hadoop cluster as the data analysis and processing server and Javaweb as the interactive application system, and highly integrates remote access, resource retrieval, shared application, data analysis and processing and other application services under the service-oriented structural framework. It effectively solves many deficiencies in the integration, circulation, application and management of online teaching resources in colleges and universities, and realizes the real co-construction and sharing of online teaching resources in colleges and universities.

2   System construction

The overall framework of the platform is divided into access layer, business application layer, data processing layer and source data layer, as shown in Figure 1. [4] Among them, the data processing layer is mainly realized by Hadoop framework under big data technology. Hadoop framework adopts cluster deployment, and four function nodes are set according to Master/Slave mode, including one NameNode, one Jobtracker and two DataNode. The hardware configuration of each node consists of 3.2GHz 4-core CPU, 16G memory and 500GB hard disk. In terms of software programs, the underlying operating system of each node in Hadoop cluster is Linux CentOS 7, and the Hadoop framework version is 2.7.5. After Hadoop is installed, the FileSystem, Replicas, Mapreduce, Yarn, HDFS and other components are set and adjusted in turn to complete the storage and management of many different types of network teaching resources. [5]
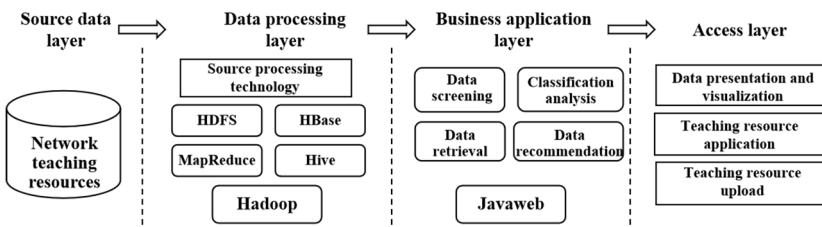

Fig. 1. The overall frame structure of the platform

In addition, the platform will complete the design and development of search engine based on Lucene algorithm. Lucene framework can build an index library according to the domains in different document objects, and supports QueryParser class to process the content input by users to obtain the Query object, and then performs the search through IndexSearch class, and returns the results matching the search conditions to the user. [6] Under the Hadoop framework, the index library will be built based on HDFS, and Lucene will run under the MapReduce computing model.

Finally, the construction of each functional application module of the platform will depend on J2EE technology system. Java is the basic development environment, MyEclipse V 2022 is the integration tool, Tomcat 8.0 is the Web server and MySQL 5.7 is the database server. Complete the configuration of Tomcat in the Preference option under MyEclipse. Then, based on the Spring architecture, the integration and encapsulation of the whole system are completed.

3   Functional implementation

3.1   Resource upload

Teacher users can log in to the platform through the client browser. The functional authority of teacher users focuses on the organization and management direction of the platform, and can upload, maintain and delete the network teaching resources according to the classification standard of the platform. When teacher users upload online teaching resources, the platform will call the WebHDFS interface to support users to operate HDFS using HTTP/HTTPS protocol, so that resource data can be stored in HDFS conveniently. [7] At the same time, the platform uses Lucene framework to segment the titles, abstracts, teachers' information, links and other contents of teaching resources to form a single index item, and at the same time uses IndexWriter to update the index database. [8]

3.2   Retrieval and application

When student users log on to the platform, they can enter keywords in the search box of the interface to complete the retrieval of online teaching resources. First, the platform will automatically segment the content, and use TF-IDF formula to calculate the similarity weight of keywords in the index database. Formula TF-IDF is shown in Formula 1, where W represents the similarity weight, TF represents the frequency of keywords appearing in the index library, IDF represents the inverse frequency of index items containing keywords in all index items, f represents the frequency of keywords appearing in the index library, $c(T)$ represents the total number of index items in the index library, n represents the total number of index libraries, and m represents the number of index libraries where keywords appear. [9] Secondly, the similarity score is calculated by cosine similarity algorithm according to the keyword weights, and the results are sorted according to the scores to complete the presentation of the final search results. As shown in Formula 2, it is a formula for calculating the cosine value, where A and B represent two text contents respectively, and S represents

the cosine value, which also refers to the similarity of the two text contents. [10] Table 1 and Table 2 show the platform simulation test results. Compared with the same type of algorithms, TF-IDF under Lucene framework has the highest consistency, the lowest confusion and the better retrieval effect when dealing with the same amount of document content as traditional Bag of Word and Mallet model.

$$W = TF * IDF \,, \quad TF = \frac{f}{c(T)} \,, \quad IDF = \log \frac{n}{m} \tag{1}$$

$$S = \cos(\theta) = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}} \tag{2}$$

Table 1. Calculation results of TF-IDF algorithm

Word	Index library occurrence times	TF	IDF	TF-IDF
Chemistry	25	0.047	1.651	0.077
Management	46	0.125	1.436	0.179

Table 2. Algorithm comparison results

Algorithm model	Consistency	Confusion degree
TF-IDF	0.69	-0.94
Bag of Word	0.61	-0.88
Mallet	0.57	-0.82

After obtaining the search results returned by the platform, student users can see the corresponding teaching resources on the page. The platform not only includes the regular course teaching resources of various professional disciplines, but also brings ideological and political, traditional culture, innovation and entrepreneurship, mental health and other contents into the teaching resource database, forming a multi-dimensional and three-dimensional network teaching resource supply service system in colleges and universities.

3.3   Data statistics

Under this function module, teacher users can make statistical analysis on the usage of the platform by students and users in a certain period of time. Common statistical analysis contents include platform usage time, course resource retrieval amount, cumulative retrieval amount, etc. Relevant data results will also be displayed in the form of charts to facilitate users' intuitive viewing and application. Figure 2 shows the fluctuation curve of the retrieval volume of curriculum resources of college students in recent one month.
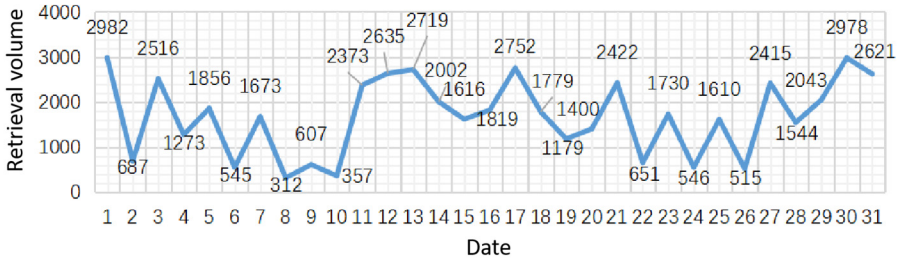
Fig. 2. The retrieval volume of curriculum resources of student users

The platform integrates D3.js data visualization class library into the front-end interactive interface, and completes the mapping relationship and layout display of system data based on SVG canvas. Among them, visual elements can be circles in scatter charts, rectangles in column charts, histograms, lines in line charts and so on. [11] The core of layout is to make clear the position coordinates of each element, and the operation needs to be completed with the help of the layout function in D3.js. The following is part of the code for D3.js to complete data traversal and add visual elements.

```
const rects = svg.selectAll('rect')
    .data(dataset)
    .join('rect')
    .attr('x', d => 20 + d * 70)
    .attr('y', 20)
    .attr('width', 50)
    .attr('height', 100)
    .attr('fill', d => colors[d % colors.length])
        return colors[d % colors.length]
    })
```

In addition, the overall operating efficiency of the platform is also evaluated accordingly. Under 3102578 original data, the experimental results of data search between Lucene framework and SQL command of traditional database under the three-dimensional network teaching resource platform are shown in Table 3. The results show that the retrieval efficiency of Lucene framework far exceeds that of traditional SQL commands, which can meet the high concurrency control of the platform and ensure the fluency of users.

Table 3. Lucene framework retrieval comparison test results

Field name	Content	SQL command		Lucene framework		Judge
		Result set	Time (seconds)	Result set	Time (seconds)	
ID	22514	1	653	1	197	Same
Name	Xi%	461	1240	461	243	Same
Msg	Systematicness %	5397	1798	5397	375	Same

4   Conclusions

In order to promote the reform of the network teaching resources supply system in colleges and universities, this paper aims at many shortcomings faced by the current network teaching resources in the practical application process, and constructs a three-dimensional network teaching resources platform. The platform effectively improves the utilization efficiency of teaching resources, optimizes the teaching process, and realizes the co-construction and sharing of network teaching resources data in colleges and universities. In the follow-up research, the platform will further enhance the abundance of online teaching resources, optimize the retrieval and recommendation algorithms of the platform, and realize the digital and intelligent transformation and upgrading of the overall education and teaching mode in colleges and universities.

References

1. Xu Shihua et al. Exploration of Practical Approaches to the Informatization Construction of Higher Education[J]. China Management Informationization.2023.01.
2. Chen Lingling et al. Investigation and Research on the Current Situation of Network Teaching Resources Construction in Domestic Universities[J]. Science & Technology Information.2023.02.
3. Yuan Shenfang. Exploration on the Construction of Network Teaching Resource Platform in Higher Vocational Colleges[J]. PC Fan.2017.02.
4. Yang Cuijie. Research on the Development and Application of Network Teaching Resource Platform[J]. Information Recording Materials.2019.04.
5. Shi Fangxia, Gao Yi. Application Analysis of Hadoop Big Data Technology[J]. Modern Electronics Technique.2021.09.
6. Hong Piao. Research and Implementation of Search Engine Based on Lucene[D]. Dalian University of Technology.2016.09.
7. Zhu Yuqing et al. Design of HDFS Content Management Module Based on Web[J]. Telecommunications Network Technology.2017.05.
8. Hu Bo. Research and Implementation of Vertical Search Engine Based on Lucene[D]. Beijing University Of Technology.2016.06.
9. Zhang Junfei. Improved TF-IDF Combined with Cosine Theorem to Calculate Similarity of Chinese Sentences[J]. Modern Computer.2017.11.
10. Jia Xiaoxia. Design of Dynamic Retrieval System of Network Resources Index Information Based on Lucene[J]. Microcomputer Applications.2021.03.
11. Quan Xin. Design and Implementation of Data Visualization System Framework Based on D3.js[D].Beijing Jiaotong University.2016.05.