Development of Ideological and Political Practice Research and Analysis System Based on Big Data Technology

Yuhong Bai(✉)

Hulunbuir University, Hulunbuir, Inner Mongolia, China
173312115@qq.com

Abstract. With the development of modern information technology, the era of big data has arrived. In this context, the ideological and political practice education work in colleges and universities has been affected to a certain extent. Therefore, how to use big data technology to vigorously develop ideological and political practice education in colleges and universities has become an urgent problem that ideological and political educators need to think about and solve. This article is based on the current situation of ideological and political practice, based on the information of learning ideological and political courses, organizing or participating in ideological and political-related activities, etc. As a means, combined with Java, Spring framework and other Web development technologies to build an ideological and political practice research and analysis system. The system not only achieves the purpose of research and analysis of ideological and political practice, but also achieves the goal of effectively improving the comprehensive quality of students, but also further promotes the development of ideological and political practice education and teaching.

Keywords: ideological and political practice · research and analysis · big data technology

1 Introduction

Ideological and political practice is an important part of ideological and political education for college students. Its purpose is to establish a style of study that combines theory with practice for college students through practical links, and to improve college students' ability to analyze and deal with practical problems using the Marxist world outlook, outlook on life and values. Students' ideological and political quality (Wei 2021). Practice is an inevitable way to achieve the purpose of ideological and political education in colleges and universities. Whether it is the ideals and beliefs of striving for the realization of the Chinese dream, or scientific theoretical knowledge and practical skills, only through the perception and application of practice can they be better absorbed by college students. Practice enriches the content and form of ideological and political education in colleges and universities. Through the social practice activities of walking into the community and taking root at the grassroots level, college students can deeply

Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 673–682, 2023.
https://doi.org/10.2991/978-94-6463-034-3_69

integrate their own cognition with reality, so as to correct their attitude towards life and establish correct value orientation and goals (Liu 2021).

However, the previous ideological and political practice has problems such as lack of systematicness and feedback, which greatly reduces the actual effect of education. In the era of big data, the goal of systematic collection and analysis of ideological and political practice information can be achieved by using big data technology, and the function of online feedback on ideological and political practice can be achieved with the help of Web technology, thereby ensuring the accuracy and comprehensiveness of data (Cai 2021).Therefore, this paper believes that the ideological and political practice research and analysis system based on big data technology and combined with Web technology can not only summarize the ideological and political practice of all college students, but also effectively analyze the ideological and political practice, and evaluation, so as to help college students guide their own practice with correct thinking, continuously improve students' ideological awareness and cognitive ability, and then promote the deeper development of ideological and political education in colleges and universities [2].

2 Overview of Related Technologies

2.1 Big Data Technology

Big data refers to the collection of data that cannot be captured, managed and processed by conventional software tools within a certain period of time. Big data has four characteristics: Volumet, Variety, Velocity, and Value, which is generally called 4V. The tools used to process big data are called big data technologies, and Fig. 1 is a diagram of a commonly used big data framework. As can be seen from the figure, data can be divided into real-time data and offline data, structured data (MYSQL, ORACLE, etc.), semi-structured data (file logs) and unstructured data (video, PPT, etc.), generally required The data analyzed are structured data and unstructured data; the big data technology system is divided into seven parts: data source layer, data transmission layer, data storage layer, resource management layer, data calculation layer, task scheduling layer and business model layer., in which the data source layer refers to the data in the business, such as customer visits, order information, etc.; the data transmission layer is mainly divided into Sqoop, Flume, Kafka, etc. according to the structure of the data; the data storage layer, the files commonly used in Hadoop The main storage component is HDFS, and HBase is non-relational data; the resource management layer, that is, the management of Yarn resources, is responsible for allocating memory and network work; the data computing layer is divided into offline scheduling and real-time scheduling. Among them, offline computing is mainly responsible for statistical tasks, and MapReduce and SparkCore memory computing are commonly used; real-time computing is mainly responsible for real-time computing with high concurrency (high concurrency means that the system can process a large number of requests in parallel at the same time), such as Double Eleven events, etc. The main tools used are Spark Streaming real-time computing and Flink. The task scheduling layer is responsible for arranging the sequence of tasks and other tasks. The following is a brief description of the components involved.

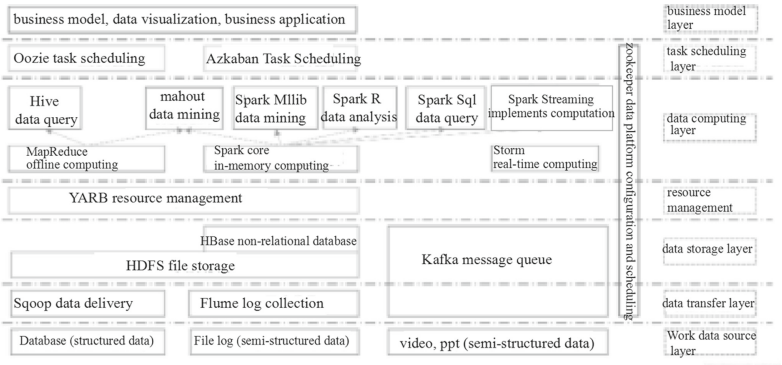

Fig. 1. Commonly used big data framework diagram (Source: Baidu Library)

Flume, used to collect unstructured data, is characterized by distributed, high reliability, high fault tolerance, and easy customization and expansion. Its collection principle is shown in Fig. 2. The working principle of Flume is to collect data from multiple clients first. Then store the data in HDFS. At the same time, the Flume data stream performs simple processing operations when processing log data, such as filtering, format conversion, etc. Sqoop is used to collect structured data, it is a bridge connecting relational database and Hadoop, that is, by using Sqoop, the data of relational database can be imported into HDFS, HBase and Hive of Hadoop system, and data can also be extracted from it, and it also has the effect of speeding up data transfer. HDFS is a distributed file system. It has good scalability, high fault tolerance and high storage characteristics. Its basic principle is to divide files into data blocks of the same size and store them on multiple machines. It is generally used for data archiving. YARN refers to the resource management system, which is responsible for the resource management and scheduling of the cluster. It has the following characteristics: good scalability and high availability; it can manage and schedule various types of applications in a unified manner. Spark and Flink are both memory-based distributed parallel computing frameworks, but the former is mainly suitable for iterative MapReduce algorithms such as data mining and machine learning. Its representation is Java Object, and the latter is mainly used for streaming data. The form of expression is logical plan.

2.2 Web Technology

The simple application of the web is that the client sends a request. After the web server receives the request, it determines whether the request is to access the extended function provided by the web server. If so, the request is processed through the written program, and the result is returned to the web server software. The web server software then returns the result to the client. The process is shown in Fig. 3. The connection between the client and the web server is to realize the request/result transmission through the extended interface provided by the server software. A more complex Web application is shown in Fig. 4, which is a client-load balancer (Nginx)-intermediate server (Node)-application server-database model. This mode is generally suitable for applications with a large

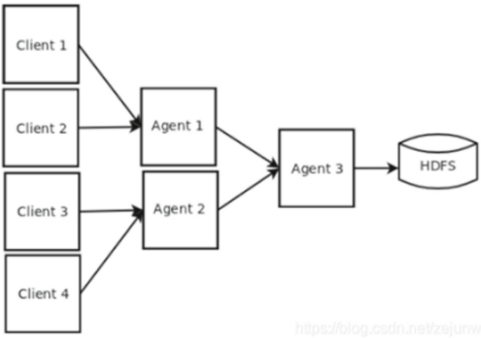

Fig. 2. Flume collection principle (Source: Baidu Library)

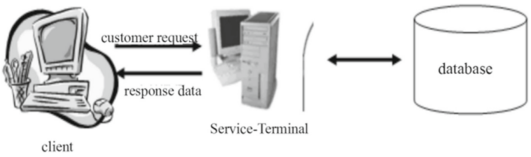

Fig. 3. B/S Architecture (Source: Baidu Library)

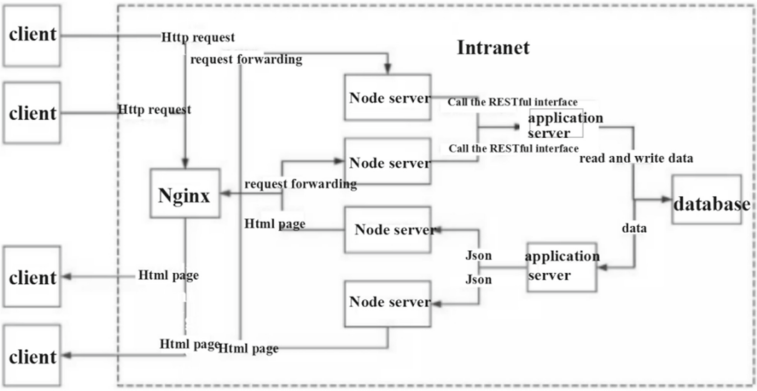

Fig. 4. Client - Load Balancer (Nginx) - Intermediate Server (Node) - Application Server - Database Mode (Source: https://www.it610.com/article/2063200.htm)

number of users and high concurrency. The basic process: the client first initiates an Http request to the load balancer, the load balancer forwards the request to the Node server cluster, the Node server first receives and parses the request, and then calls the RESTFUL interface exposed by the application server. For interaction, after the interaction, the data is sent back to the application server, and the json data is returned to Node. The Node layer renders the template + data combination into html and returns it to the reverse proxy server until it is returned to the client.

Commonly used WEB development technology can be divided into static WEB development technology and dynamic WEB development technology.

Static WEB development technology involves HTML technology and XML technology. HTML technology, its files are <HTML>, </HTML> as the beginning and end. <head>…</head> is the header information of the file, the information located in the middle of <title>…</title> will be displayed on the page; between <body>…</body> The code is the body of the HTML file, and the content is displayed in the client's browser. HTML is the basis for making web pages, and sometimes js, css, jquery and other technologies are used. XML technology is a markup language, which can achieve barrier-free effect during data transmission and display the data on the user's browser.

Dynamic WEB development technology, generally used technologies are ASP technology, ASP.NET technology, PHP technology and JSP technology. The following is a brief description of the technology: ASP technology, the abbreviation of Active server page, is defined as a dynamic web page. Its formation needs to rely on HTML, javaScript and CGI technology. Compared with other technologies, its advantages are running efficiency and flexible programming. ASP.NET technology, its advantage is that by using the.NET framework, the calling interface is more convenient; DLL files are generated at compile time, which improves the running speed. PHP technology, which is characterized by the use of free and open source code. JSP, is a simplified Servlet.

2.3 Development Environment

According to the above technical requirements, complete the construction and deployment of the development environment. The development environment of the system is divided into two parts, one is the construction of the big data environment, and the other is the development of web applications.

The requirements for building a big data environment are as follows: The big data system architecture is based on the Linux system, and Hadoop will be deployed in a cluster, with a total of 5 sets, one as the master node and the rest as data nodes. At the same time, the deployment of functional components such as Flume, Sqoop, HDFS, Spark and MapReduce is completed under the Hadoop framework.

The requirements for developing Web applications are as follows: The system as a whole is to use Web technology to complete the development of the system and publish it to the Web side. The specific requirements are as follows: HTML and Javascript are development languages for client-side technologies. On the server side, the Tomcat server, Java technology development language and spring framework are used to complete the construction of the Web server side. The database server selects MySQL. Among them, spring completes its configuration by installing JDK, creating a Spring project, creating a maven project, and importing Spring's jar package. The involved bean creation code is shown in Fig. 5. The code includes bean definition, implementation and xml configuration. Among them <!-- id = "unique identifier" class = "implementation class permissions"-->, <!-- lazy-init:true tells the Spring container, no longer instantiate the corresponding bean when the container starts -->.

Through the introduction of the above key technologies, the overall environment for system development, the configuration of related software and tools are determined, and

```
package com.xk.spring.kp03_container.s1_basic;public interface IContainer {
    public void testcontainer();
}
package com.xk.spring.kp03_container.s1_basic;
public class ContainerImlp implements IContainer {
    public ContainerImlp() {
        System.out.println("ContainerImlp.ContainerImlp()");
    }
    public void init() {
        System.out.println("ContainerImlp.init()");
    }
    public void destroy() {
        System.out.println("ContainerImlp.destory()");
    }
    public void testcontainer() {
        System.out.println("------------------------");
    }
}
<?xml version="1.0" encoding="UTF-8"?>
<beans xmlns="http://www.springframework.org/schema/beans"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.springframework.org/schema/beans
http://www.springframework.org/schema/beans/spring-beans.xsd">
    <bean id="testContainerBean"
class="com.xk.spring.kp03_container.s1_basic.ContainerImlp"
    init-method="init" destroy-method="destroy"/>
</beans>
```

Fig. 5. Create bean code (Original)

the technical feasibility of the ideological and political practice research and analysis system as a whole is also clarified.

3 System Requirements

3.1 System Requirements

The ideological and political practice research and analysis system is oriented to the entire ideological and political education major in colleges and universities, and aims to change the current difficulties faced by colleges and universities in the practice of ideological and political practice, such as large amount of data and information, and cluttered content [4]. In the process of designing the system, the system needs to consider the practicability and efficiency of the system function in the selection of data processing tools. In addition, the system also needs to meet the relevant work requirements of teachers and workers to analyze data. Therefore, the overall goal of this system is to take students' ideological and political practice data as the research object, establish a research and analysis system that supports teachers to explore and analyze students' ideological and political practice, and make it easier for teachers to study students' ideological and political practice activities and discover existing problems. Provide a comprehensive platform for making reasonable suggestions. Referring to the relevant research and analysis systems and the basic goals of this system development, the system requirements mainly include: providing data processing functions to ensure the normal operation of data; supporting data cleaning to create clear visual effects; providing online evaluation functions to achieve the purpose of the quick test [7].

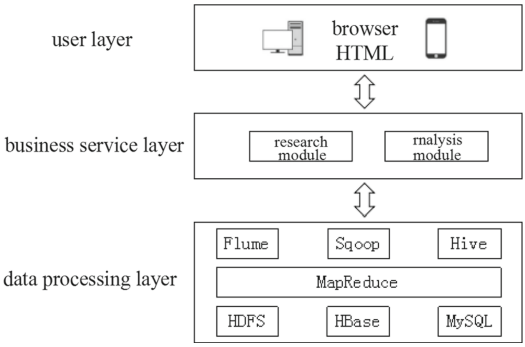

Fig. 6. The overall framework of the system (Original)

3.2 Overall Design

The overall design of the ideological and political practice research and analysis system is based on the concept of layers. This idea is not only conducive to the overall function and technical planning of the system, but also to providing strong technical support for practical education in colleges and universities [3]. The system adopts a B/S architecture, as shown in Fig. 6. At the user layer, teachers can log in to the system through computers and mobile phones, and enter the business service layer. Teachers can not only view all the ideological and political practice data collected through research, but also use the analysis module to obtain valuable information from students in ideological and political practice activities. information. The design of these two modules is not only helpful for teachers to discover the problems existing in students' ideological and political practice activities, but also helps teachers to give reasonable suggestions [5]. At the same time, it can be seen from the side that ideological and political practice is highly valued by students and teachers. The normal operation of the business service layer in the system needs to rely on the data processing layer, because the data processing layer not only needs to perform the responsibility of receiving requests from the business service layer, but also needs to provide the business service layer with the service of query result feedback. The query result is finally fed back to the user layer by the business service layer, thereby ensuring that the request sent by the teacher is directly converted into the query result. The tools involved in the data processing layer include Flume, Sqoop, Hive, MapReduce, HDFS, HBase, and MySQL. The design of these tools is conducive to the efficient operation of the system.

4 Function Realization

Teachers can directly enter the ideological and political practice research and analysis system by logging in through the browser. There are two modules in the interface of the system home page, namely the research module and the analysis module. Among them, the research module summarizes various ideological and political practice activities of college students, and the analysis module analyzes and evaluates the content in the research module.

4.1 Research Module

Research is a basic and long-term work. It must have a set of feasible operating mechanisms, and make detailed plans for the objectives, tasks and content requirements of the research work [1]. According to the theme and needs of this paper, the research process is designed as follows: First, the purpose of the research is determined to provide a basis for teachers' scientific decision-making. Secondly, the necessity of the investigation is determined according to the purpose of the investigation, that is, the investigation is not only a fundamental method to understand the students' ideological dynamics, but also an effective way to carry out ideological and political education. Third, determine the development goals of the research, the goal is to expand the scope of students' ideological and political practice and stimulate more students to carry out ideological and political practice activities. Finally, gather data. The collected data includes three parts: one is the information on ideological and political learning, such as the number of studies, depth, etc.; the second is to organize and participate in ideological and political-related activities, such as conferences, exhibitions, discussions, community assistance, and speeches; the third is for students The scores of the ideological and political examinations, through the statistics of these three parts of the data, then select the students' ideological and political practice, and then evaluate the three grades of students who are excellent, qualified, and yet to be improved. There are two ways to collect data. One is to use the big data tool Flume to collect log file data from web pages and store them in HDFS; the other is to use Sqoop to directly store data in the data repository of college students' ideological and political learning. In HDFS; the second is to collect data through questionnaires. When designing the questionnaires, WeChat's answering applet and Java language are used. Finally, the module realizes the function of collecting data, and the data there shows the characteristics of large amount and comprehensiveness.

4.2 Analysis Module

The analysis module is the analysis of the data collected by the research module. Its functions are: first, it can process a large amount of data at the same time; second, it can deeply clean the data, so that the data presents a neat effect; third, it can be evaluated online, which is helpful for teachers to provide targeted guidance to students. The analysis module is divided into three parts. First, clean the data. The objects of data cleaning are missing values, format content, and logical errors. For missing values, the cleaning methods of deletion, mean filling and hot card filling can be adopted; the processing of outliers can be realized by statistical analysis, model detection and the 3α principle; the processing of noise can be solved only by applying the binning method and the regression method. Second, analyze the data. The analysis process is as follows: First, use Hive to analyze the cleaned data, secondly use MapReduce to calculate the data, and store it in HDFS, and finally store the data in HDFS in the database, if the data is unstructured data. Data storage needs to be done with the help of HBase. The design of the first and second parts greatly improves the speed of processing data and improves the efficiency of analyzing data. Third, evaluation data. Teachers can evaluate students' ideological and political practice activities online. The implementation process is as follows: extract evaluation data from the database, for example, select a certain class

of a certain major in a certain university in the database in 2021. The relevant data of ideological and political practice in the last semester of the year are used for evaluation. For example, the evaluation method is to leave a message, and you can set the top message; at the same time, students can also leave a message for the activity. This part not only plays the role of result feedback, but also achieves the effect of interaction [6].

5 Conclusions

The ideological and political practice research and analysis system is based on big data technology and combined with Web technology to realize the functions of ideological and political practice to store large amounts of data, organize data and analyze data. The system as a whole takes Java Web application as the core. Through the joint application of client browser, Web server and big data architecture, it can help teachers to achieve rapid query and analysis of ideological and political practice data information, breaking through the current ideological and political practice that cannot achieve data collection, the technical barriers of analysis, improve the flow and interaction of data and information, help teachers to quickly obtain the connection between data and information from multiple aspects, establish an ideological and political practice research and analysis system, and provide for further promotion of deeper ideological and political practice education.

Acknowledgements. The 2020 Hulunbuir College Scientific Research Project "Research on the Cultivation Mechanism of University Network Cultural Works in Major Epidemic Prevention and Control-Taking Hulunbuir College as an Example", Item Number: 2020SZC10.

References

1. Chen Bihua, Huang Huali. (2001). The role of research work in ideological and political education in colleges and universities. J. Journal of South China Normal University (Social Science Edition).
2. Jin Rong, Liu Wei, Li Chen. (2021). Construction and implementation of practical courses in ideological and political theory courses in colleges and universities. J. Journal of Taiyuan City Vocational and Technical College.
3. Huang Weihong. (2021). Exploring the innovation of ideological and political teaching in five-year higher vocational ideological and political courses in the era of big data. J. Journal of Jiamusi Occupational College.
4. Shen Feiwei, Chen Xiaoling. (2021). The practical logic and optimization strategy of the platform of "learning to strengthen the country" to improve the teaching quality of ideological and political courses in colleges and universities. J. Shenzhen Social Sciences.
5. Zhang Haidong Li Qianghua. (2020). Research on the construction of practical teaching resource sharing platform for ideological and political courses in colleges and universities. J. Journal of Luliang University.

6. Zhang Hongmei. (2013). On the Necessity of the Construction of Quality Education Practice Bases for College Students—Based on the Practical Reform of Ideological and Political Theory Courses in Higher Vocational Education. J. Teaching Practice Research.
7. Zhang Nan. (2021). Exploration of Informatization of Ideological and Political Education in Colleges and Universities in the Age of Big Data—Comment on "Theory and Practice of Flipped Classroom for Ideological and Political Courses in Colleges and Universities in the Internet Era". J. Research on Science and Technology Management.