

# Application of Big Data Technology in College Information Construction

Wei Deng

Shanxi Normal University

**Abstract.** As a new technological means, big data plays an important role in the construction of intelligent campus in Colleges and universities. Taking Shaanxi Normal University as an example, this paper introduces the development of college information construction, and elaborates on the typical application scenarios of big data in colleges.

## 1 Background

With the rapid development of information construction in colleges and universities, the application system of various business departments in colleges and universities has accumulated a great deal of data. This lays the foundation for the application of big data technology. Big data [1] not only refers to massive data itself, but also includes a series of technical methods for data storage, processing, analysis and visualization. It is an effective means to explore its hidden value from massive data.

The information construction of colleges and universities generally goes through three stages: infrastructure construction [2], digital campus construction [3] and intelligent campus construction [4]. In the stage of infrastructure construction, hardware is the main construction content. It includes network equipment, one-card system, access control system. It has accumulated plenty of students online authentication data, campus card consumption, access control card data. In the stage of digital campus construction, software is the main construction content. It mainly includes business management systems such as scientific research, educational affairs, assets, and personnel. Its purpose is to support the daily operation of various business in colleges and universities, transform offline work into online services, improve working efficiency, and improve users' experience. Smart campus is the construction goal of colleges and universities at present. It integrates advanced technologies such as Internet of Things, big data and artificial intelligence to improve service quality, enrich teaching means and improve decision-making methods.

## 2 Present situation of construction

The network information center of Shaanxi Normal University began to explore and try the construction of big data in colleges and universities in the second half of 2014. The construction content includes two parts: big data platform and analysis modeling system of business

data. The construction ideas mainly focus on the collection, storage, analysis and visualization of structured and semi-structured data generated by various business systems throughout the school. The purpose of the construction is to provide objective data support for the school's scientific research, teaching, management, and services.

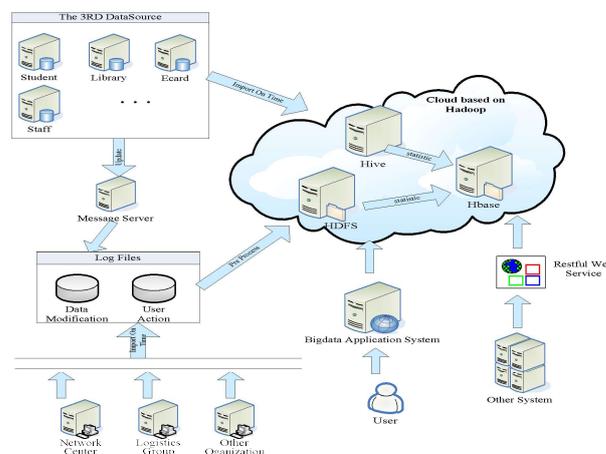


Figure 1. Platform Architecture

In the construction of big data platform, a big data management platform is built with Cloudera 5.9 (open source system) on 20 physical servers, with a total storage capacity of 300TB (Figure 1). On this basis, Sqoop and Linux script are used to collect data and HDFS and Hbase are used to store data. Besides, Mpp cluster and Redis cluster are used to store real-time data and hotspot data. MapReduce and Spark are used to analyze data while Spring MVC and Echarts provide a visual representation of the data. The big data platform has access to data for a total of 12 business systems. This includes personnel system, scientific research system, educational administration system, graduate management system, employment system, one-card system, book lending system, library access control system, dormitory access control system, wireless network management system,

network authentication system, online behavior management system. At present, the platform supports the division of authority according to five roles: system administrator, school administrator, department/college administrator, instructor, teacher and student. The data and statistical analysis results stored in the platform can be used for external data sharing services in the form of API.

### 3 Typical application

The analysis modeling system of business data realizes decision support module for business departments and personal data center for teachers and students of the whole school. Moreover, it realizes the early warning module for student management and the heat map of campus wireless network for network management.

1. Decision support module establishes models by setting analysis indexes for different business systems. It uses a variety of statistical reports to visualize the model. The most commonly used report components are bar charts, line charts, and scatter charts. The bar chart is mainly used for the horizontal comparison of indexes in the model between different colleges. The line chart is mainly used as a vertical comparison of the indexes in the model at different time points (ie, showing the trend of development). The scatter chart is mainly used to analyze the correlation between the two indexes when modeling across business systems. Currently, the established models mainly include:

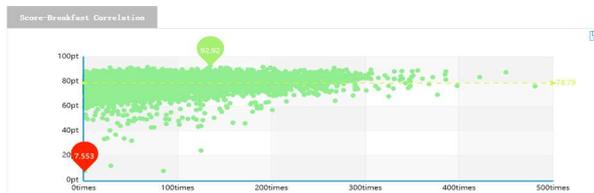


Figure 2. Breakfast frequency-score

Correlation analysis of students' breakfast frequency and test scores (Figure 2);

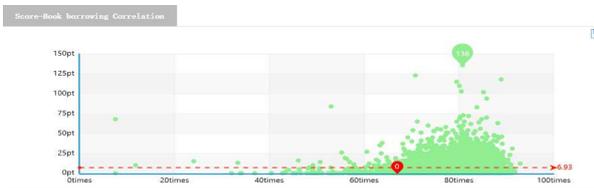


Figure 3. Book borrowing-score

Correlation analysis of the times of student's book borrowing and test scores (Figure 3);



Figure 4. Freshmen average score trend

Quality analysis of undergraduate freshmen (Figure 4);

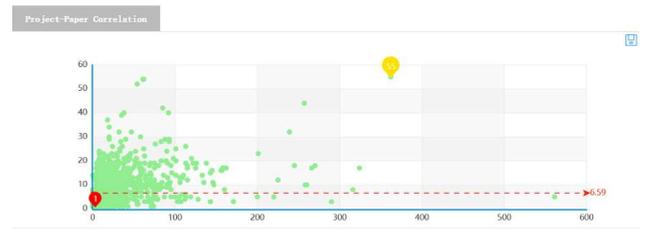


Figure 5. Project-paper

Correlation analysis of the number of teachers' projects and papers (Figure 5);

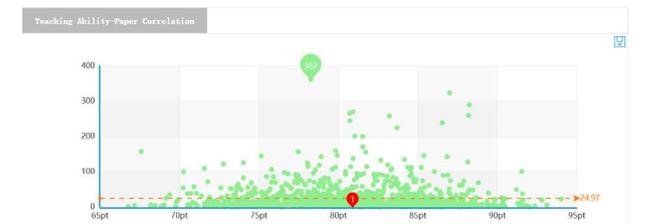


Figure 6. Teaching ability-paper

Correlation analysis of teachers' teaching ability and the number of papers (Figure 6);



Figure 7. Scientific research trend

Development trend analysis of scientific research achievements of the whole university and various colleges (Figure 7);

Comparison of the number of scientific research achievements between different colleges;

Automatic generation of quality analysis report of the annual employment;

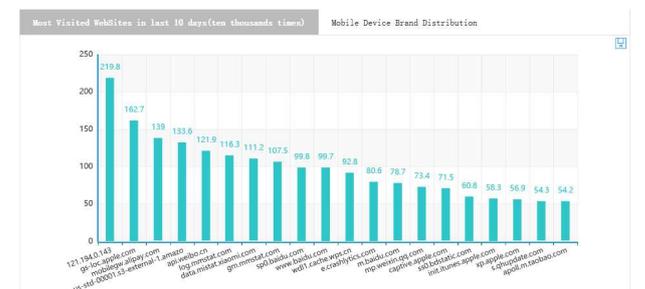


Figure 8. Popular website

Analysis of popular website of on-campus visits (Figure 8);

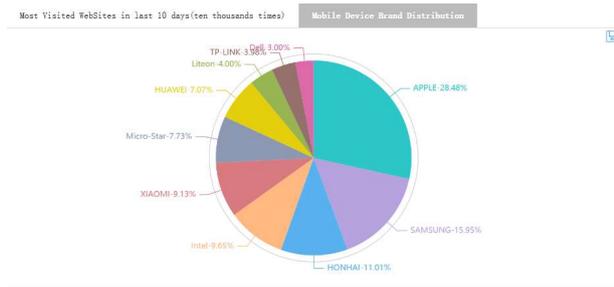


Figure 9. Brand distribution

Distribution of mobile devices' brand used by staffs and students (Figure 9).

Combined statistical analysis of the attributes of basic information of teachers (Basic information attributes of teachers include: Native place, college, gender, professional title, education background, age, category, time of admission);

Combined statistical analysis of the attributes of basic information of students (Basic information attributes of students include: Native place, college, gender, nationality, enrollment time);

2. The personal data center is a page that aggregates the basic information of the teacher or student and various types of school data and displays the portraits. Teachers and students do not need to log in to each business system separately to view separately.

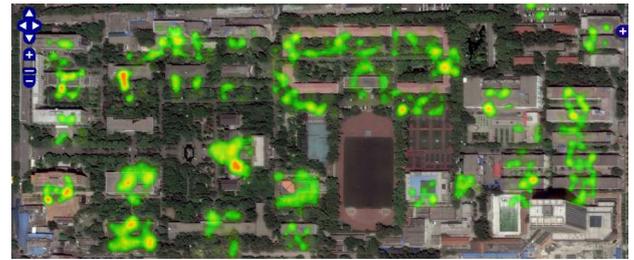
After logging in, teachers can see their personal basic information, course setting, scientific research results, one-card consumption, device online behavior and other data. In addition, teachers can also see portraits based on their teaching ability, scientific research ability, consumption ability and awards.

After logging in, students can see their personal basic information, course arrangement, test results, one-card consumption, online behavior of devices, library borrowing and other data. Moreover, they can also see portraits according to their learning ability, consumption ability, time spent online, social activity ability and award.

3. Early warning of students is an attempt in student management. At present, it mainly makes early warning for poor students and suspected non-return early warning, which can provide a basis for the grant and student safety management.



Changan Campus



Yanta Campus

Figure 10. Wireless network heatmap

4. Campus wireless network heat map (Figure 10) is a real-time heat map of campus mobile devices distributed by correlating the data of wireless network management system and authentication system. It can show in real time the degree of people's density distributed in the school, and what equipment is available in every building, floor and room on the Internet. Through further analysis of the heat map, it is possible to diagnose whether the distribution of wireless hot-spots in schools is reasonable, to grasp the trajectory of each student and the places that they often go, for monitoring the safety of students who are focused on.

## 4 Conclusion

Big data technology can effectively integrate the data generated by the existing business systems of colleges and universities within the scope of the whole university. Through data correlation analysis, we can mine the core value of data, provide deep information value-added services for college teachers and students, and provide data support for scientific management and decision-making in colleges and universities. Finally, we will transform the data into school benefits, improve management efficiency, promote the quality of scientific research, teaching and service, and enhance the comprehensive competitiveness of schools.

## References

1. Wu Yakun, Guo Haixu, Wang Xiaoming. Research Overview on Big Data Technology I [J]. Journal of Liaoning University (Natural Sciences Edition), 2015, 03 (8): 236-242.
2. CAO Chuang. The Present Situation and Countermeasures of the Development of Management Information System in Colleges and Universities [J]. Education Teaching Forum, 2018, 47 (4): 12-13.
3. Tand Kailin. A Brief Analysis of Digital Campus Construction and Application [J]. China Computer & Communication, 2018, 20 (4): 237-238.
4. Yang Long. Research on university informationization service based on big data and artificial intelligence [J]. Experimental Technology and Management, 2018, 35 (11): 153-156.