Research on Construction Method of Regional Education Data Center in the Era of Big Data

Hai-Feng Li* and Hong-Rui SUN
Dalian Jiao Tong University, Dalian, Liaoning, China
*Corresponding author

Keywords: Big Data Education, Regional Education Data Center

Abstract. Based on the analysis on the present situation of China's current education, this paper lists some problems of the traditional education and data center. It analyses the method of constructing the regional education data center and how to integrate education with its ecological system, and initially plans and designs regional education based on big data, and recommends to construct the regional big data education ecological system. It provides ideas for the transformation of education in China and provides references for the construction of regional education data centers in the era of big data in China.

1 Introduction

Over 2013, the world entered the era of big data. The use of professional data processing tools can extract valuable information to promote the development of all walks of life, so is it for education industry. How to extract valuable data from mass education data and analyze them is an important step in education development of our country. On the other hand, people have higher and higher requirements for learning efficiency and quality. Traditional education cannot meet people's needs, and technological progress and learning demand jointly promote the development of big data education. Therefore, it is indispensable to build a big data education management institution—regional education data center.

In the era of big data, the development of data processing technology has made data a resource. Through the establishment of the data center, we can propose more scientific, accurate and efficient solution for regional education, which provide the guidance on many problems, such as the information island phenomenon, uneven distribution of education resources or the educational data inconsistent standards. On the other hand, the promotion of regional education data center construction theory based on big data is beneficial to the construction of China's regional education data center, which provides new development ideas for the transformation of education in China and improves the level of education.

2 Analysis of Regional Education Data Center

2.1 Analysis of the Current Situation of Basic Education

According to the data from IRESEARCH (as shown in figure 1 below), the admission rate of K12 (basic education) in China is only about half of the enrolled population at the initial stage of growth. Less than half of high school students have successfully progressed to undergraduate. In other words, only about 25% of the total number of students can successfully enter the undergraduate level and receive higher education. By contrast, countries such as the US, Japan and Germany are at least 10 percent higher.
At present, most areas of our country are still keeping the traditional education model. At the same time, due to the limitations of technology, capital and concept, big data and artificial intelligence education are still in the trial stage, which cannot achieve large-scale popularization in the short term. However, the traditional teaching and the development of the times are becoming increasingly mismatched. In the current social context, traditional teaching has not been able to cultivate talents that meet the needs of the development of the times, and society has increasingly higher requirements for education management departments, schools, and teachers. There has been increasing criticism of traditional unidirectional, closed, and experience-driven teaching. According to the report released by IRESEARCH, about 60% of parents do not think that elementary and secondary education can maximize the strengths of students, and more than 50% of parents believe that there are many problems existing in schools, such as solidified assessment standards, unscientific allocation of educational resources, and imbalanced teacher structure, which are shown in Figure 2.

![Figure 1. Admission Rate of K12 Education in China.](image)

**2.2 Problems in Traditional Education Data Center**

**2.2.1 "Serious Information Island Phenomenon" in Various Teaching Institutions**

Because most schools do not have a reliable Internet communication system, each school acts independently and educational data is basically one-way flow and spatial transfer, there is no horizontal data flow between schools, which will cause problems in student transfer, teacher transfer, school admission, increasing the cost of teaching time invisibly.

**2.2.2. Data Collection Equipment is Simple and Lack of Analysis Tools**

At present, school data collection only relies on electronic computers and dictation, lacking
electronic terminals, which cannot meet the amount of data required for big data analysis. At the same time, students still focus on paper textbooks, while digital textbooks are only available in some schools with economic advantages and abundant educational resources, which further restricts access to educational big data. On the other hand, lack of professional big data analysis tools, it is difficult for schools to carry out data mining, analysis, and feedback.

2.2.3. Educational Data Standards Vary, and Data Universality is Poor
Since there is no unified data technology standard, each school and other public education institutions provide services via their own technical tools, making the data universality worse. To some extent, that leads to the confusion of teaching information, making it difficult for higher education authorities to accurately obtain education data in the jurisdiction.

2.2.4. IT Equipment Updates Slowly and Cannot Keep Up with the Speed of Network Development
Schools in most areas of China are unable to meet the requirements of mass data processing in the era of big data. The pace of change has accelerated, and some schools have difficulty in upgrading their IT equipment because of the constraints of capital, technology and talent.

2.2.5. Information Security is not in Place
In recent years, there have been frequent reports of student information leaks and student information being used as goods for illegal trading. On one hand, the information protection measures are not in place and the level of security protection is not high, the security awareness of managers is insufficient. On the other hand, In terms of architecture, structure, and technology, information processing systems can no longer meet the security requirements in the era of big data networks. To strengthen data information security, we must upgrade and improve the system architecture and other aspects.

3. Establishment and Integration of Regional Education Data Center Platform

3.1 Establishment of Regional Education Data Center Based on Big Data

3.1.1 Establishment of Regional Education Data Centers at the Macro Level
The data center will adopt the three-level structure of prefecture-city cloud data center—district/county data center—teaching point for system construction. Considering the economic and technological feasibility and standardization construction, each level of institutions will have different rights and responsibilities. Specifically, due to the limited technical and economic strength of teaching sites, they are mainly responsible for the collection of educational data, the main body is students. The district/county data center is mainly responsible for collecting the data of each teaching point, and the corresponding subject is teaching point. The prefecture-level big data center analyzes and backs up the district / county date, and the data is stored in the private cloud. Teachers and students of the district / county or school can obtain the data information with corresponding permissions through registered accounts.

3.1.2 System Architecture Design at the Micro Level
The structure is divided into three layers, including data acquisition module, data analysis module and service module:

In the Data acquisition layer, the dominant data source are mainly digital teaching material, classroom speech, homework and test scores, The implicit sources are behavior data generated by operating on some learning forums or web pages, such as web page clicks, videos Browsing, forum speeches, etc., these data contain information about students' personal preferences, emotional states, etc., which is helpful for teachers to understand the state of learning and to intervene appropriately.

The data processing module is the core module of the system, which is mainly used to classify and structure the information collected in the distributed system. It mainly includes the steps of data pre-processing, data mining, data security, data storage and data modeling. Among them, data
preprocessing includes data cleaning, data integration, data transformation, etc., which can standardize data and improve the efficiency and accuracy of data analysis.

![Diagram of Regional Education Big Data Center](image)

Figure 3. Basic Framework of Regional Education Big Data Center.

Data mining and data storage mainly analyze and extract massive information data by big data technology, and then store it in distributed data warehouse for further analysis, so as to formulate scientific, refined and personalized learning plans.

Data security is essential under any circumstances. The system should be composed of several security technologies together to constitute a security system of early-warning, in-process repair and post-filing. The system firewall must be established to define the relevant access rules to ensure that the internal data of the system is not illegally accessed and outflow, at the same time, encrypting the system data to ensure the safe transmission of data, defining user verification and user rights, setting up system logs to monitor and record the operations of relevant users.

Data modeling is to integrate all kinds of fragmentary information related to users through big data technology, so as to establish user portraits for each specific group of users and subdivide the huge group into groups with common characteristics. In this way, exquisite, suitable and excellent educational products or services can be launched to improve user satisfaction and stickiness.

Service layer is divided into data sharing, data visualization, data coordination, custom service, bidding, etc., data sharing is to provide an account for each school subject, using this account to log in to the system, you can legally obtain education information of other schools, it is helpful to achieve seamless communication between schools in educational resources and abolish the phenomenon of information island, narrow education gap between school area.

Data visualization is a visual education report based on data from various aspects of education, which can provide simple, clear and accurate data reports for schools and students in its jurisdiction, for the reference of schools, teachers, students and education authorities, which promotes the two-way flow of educational information and provides some references to the institution of education, schools and teachers.

The bidding is that the data center formulates the demand for educational services on the system bidding board to guide professional third-party educational service institutions to provide educational...
service for students or teachers. This method not only saves the economy and time cost for the school to the utmost, but also provides education consultation and service for the school efficiently.

Data coordination requires data synchronization between subsystems. Within its authority, an account can access various subsystems, such as the school's financial system, educational administration system and personnel system. Data should be kept in real-time information synchronization.

In terms of customized services, teachers and students can request the big data center to test them, so as to obtain reports and customize personal education services, which can guide teachers or students to connect with third-party education service institutions and receive customized education.

Reminder, supervision, and testing mainly serve as an intelligent agent. Based on the data, the report provides necessary advance notification, supervision, and prediction after the event. For example, intelligent agents replace teachers to test students. Teachers only need to store the test questions in the examination paper database. Invigilators, corrections, and judgements can be completed by artificial intelligence. Given the development of artificial intelligence, the function will become more perfect, and automatically handle part of the teaching work for teachers.

The forum serves as an open platform for all teachers, students and families. Teachers can post courseware, assignments, and lesson plans in this forum. Students, parents, and teachers can discuss, and express opinions on teaching issues, so that all parties can join forces to participate in all processes of education.

3.2 Suggestions on the Integration of the Education Ecosystem with the Regional Education Data Center as the Core

Based on the construction of the above-mentioned regional education data center, we believe that the education data center is only a part of the regional big data education system, and relying solely on the education data center as a platform is difficult to meet the requirements of the big data education service system. The integration of the data center and the surrounding environment requires necessary improvements from the school's education facilities, talent team, and management level to achieve the organic integration of the entire education ecosystem.

3.2.1 Enhance the Level of Regional Education Informatization

Big data education is developing towards artificial intelligence. Traditional multimedia classrooms can no longer meet the requirements of intelligent education. The procurement, management, and maintenance of relevant IT equipment in each school must be procured by the regional data center. This will improve and upgrade the relevant equipment and enhance the level of school informatization.

3.2.2 Improve Technical Standards

Regional education data centers should cooperate with big data technology-related enterprises to establish technical standards and data docking, provide technical guidance for various schools, standardize the development of big data education activities in their jurisdictions, and launch new education products, education technologies, and education models, education facilities, etc., which is beneficial to form an industrial chain based on big data.

3.2.3 Accelerate the Construction of a Talent Team

A big data education center should not rely too much on data. What is more important is that it should have a team, which can make correct strategies by analyzing these data. This team is divided into three fields: business experts, big data education experts and technical experts.

Business experts mainly come from teaching institutions. Their responsibilities are to conduct field surveys, visits, and analyze the education status, summarize the education needs of the region, and plan the pre-construction and post-improvement plans of the education big data platform, continuously put forward feasibility suggestions for the development of regional data centers, so as to promote the integrity of data center functions, services and systems.

Education experts mainly come from knowledge groups such as backbone teachers. Based on
long-term teaching experience, education experts can combine new big data education models with traditional education models, which can achieve offline applications, promote the concept of big data education and provide ideas, suggestions and solutions for the development of regional education big data. At the same time, based on their own teaching summary, necessary improvements are made to the educational needs raised by business experts to make their construction plans more suitable for the actual teaching needs.

Technical experts are mainly technical personnel of major data network companies, and their main job responsibilities are to make education experts’ ideas, suggestions, and feasible solutions for the development of big data education come true, building a data center system that meets the needs of the region and keeping improving it. At the same time, in the process of system operation, experts regularly analyze users’ data through technical means, which is used as a reference and to joint education service third parties to create new education products and extend education services.

3.2.4 Change the Concept of Education Management

Under the mode of big data education, education management should also be transformed. Whether it is the macro-level education management system and structure, or micro-level such as end-of-term assessment, teaching methods, college entrance system, and title system, etc., necessary reforms must be made to adapt to requirements of precision, refinement and humanity in the data age. For example, schools should set up specialized functional departments to manage education data. In addition, it is necessary to strengthen the connection between school education and family education, attach importance to the role of family education.

3.2.5 Improve the Relevant Guarantee and Supervision System

Government supervision ensures that the system is not used illegally. If any system does not have external supervision, it will be difficult to ensure the security of system data. The government should formulate and improve relevant laws and regulations to ensure the security of educational data. At the same time, the government should increase publicity to create a social atmosphere that respects data security and enables people to consciously maintain the legitimate use of data.

4. Summary and Outlook

At present, big data education is in its infancy. This paper designs the construction method of regional education data center, which provides a reference for the construction of regional education informatization and education ecosystem. With the development of advanced technologies, such as artificial intelligence, deep learning, VR and AR, and people’s pursuit of personalized learning, the functions of regional education data centers will continue to improve and develop, and the future education model has infinite possibilities!

Acknowledgment

This research was financially supported by Liaoning Natural Science Foundation, 201800177. The subject of educational science planning in Liaoning Province, JG16DB054. Natural Science Fund of Education Department of Liaoning Province, JDL2019027

References


[5] Notice of the general office of the ministry of education on printing and distributing the key points of education informatization work in 2014 [OL].

