

# The Research of School Archives Information Resources Servicing College Education in Big Data Bases Era

Liyong Cui<sup>1, a</sup>, Li Zhang<sup>1</sup>, Qun Li<sup>1</sup>

<sup>1</sup>Jilin Agricultural University, Changchun, Jilin Province, China

<sup>a</sup>cuiliyong07@126.com

**Keywords:** Big data era, schools, profile information, resource services, higher education and research

**Abstract.** Today's society has entered the era of big data, with the computerized management of information technology to accelerate the pace of construction of infrastructure archives imperative. Combined School archival information resources management practices set forth archival information resources and services in higher education model, and the connotation of college archives information management is analyzed in its favorable conditions for basic research school era of big data archival information resources on how to serve the higher education.

## Introduction

Since the 1950 s, with the rapid development of computer, network and communication technology, human society started from the industrial society to information society. Especially in the 1990 s, the emergence of the Internet of geometric information growth. International data corporation (IDC) in 2011 launched a new digital universe study - "to extract value from the chaos", the report showed the global amount of information will increase 1 times every two years. In 2011 alone, the world is created and the amount of data replication are 1.82 ZB, compared with the same period in 2010 rose more than 1 ZB, by 2020 that figure is expected to increase to 35 ZB. In the face of information explosion, people adopted a variety of coping strategies [1]. In recent years, large data method and thought has been more and more attention and application.

In the IT industry, the term "big data" although there is no unified definition, but for its features reached a consensus [1]. Generally think that big data has the characteristics of four aspects: one is the data, i.e. data contains all of the samples, the data may be massive, may also be small; The second is the data type is various, including web logs, text, images, audio and video, the geographical position information, etc. [1]. Three is the value of low density, the unit data information content is limited; Fourth, processing speed is fast. Human society is about to or already marched into the era of big data, it will deeply affect every aspect of society.

## Big data era school archives information integration challenge

Research institutions to the big data is defined as "Gartner", big data is need to deal with the new model can have better decision-making, insight found mass force and the process optimization ability, high rate of growth and diversification of information assets. The strategic significance of big data is not the master of the large data information, but these contain meaningful data are specialized processing [2]. Schools in the face of the arrival of the era of big data, how to school enrollment, employment, education, management, and other various practical activities of great value to big data integration and inheritance of data management, and use the large data processing technology for the information mining, sorting, analysis, and to provide decision support for the development of the school, is the focus of the next school information development work.

In recent years, from such aspects as national, regional to school are to speed up the pace of information construction, but the school data integration work is still faced with many problems and challenges.

**Weak data assets.** Data as an important school of intangible assets and strategic resources, the use of the level of high and low can direct response to a school for teaching and scientific research level, management ability, and is the effective means to enhance the competitiveness of the school [2]. And still exist inside the school attaches great importance to the construction of information system, data construction consciousness weak, for schools failed to realize that data is also valuable assets of the school, will play an important role in school teaching and scientific research management, lead to internal business school system is numerous, but there is no unified management system, data and existing management chaos, inconsistent phenomenon, valuable data assets cannot play its important role. In the face of the impact of the large data, requires schools to internal constantly adjust the thought pattern, from top to bottom attaches great importance to the data on thought and action.

**Data lack of normative and unified management.** After a long time of information construction, the school formed multiple internal business systems, but the system level, such as in the application or data are separated from each other, respectively adopted different architecture technology and its development [2]. Various business system data maintenance and management work at the same time the lack of unified management and lead to data inconsistencies between different departments, difference, moreover two department released figures appear even conflict. So to determine the uniqueness of the data and the authorities, specification of the same data format and data cleaning and finishing work is the largest school workload in the data center consolidation process is also the most important content.

**Data management security is not high.** Independent schools business system are widespread, unable to realize the sharing and use of application system, application system data security servers or storage, there is no unified planning, storage security cannot be guaranteed [2]. Some schools have a data center, but lack of security management in the management of data center, the general lack of dense, classified data and general data is not to discriminate, encrypted data to do safety measures of protection. So in the process of data integration in school, how to strengthen the security of the data management is an important problem to be solved.

### **The advantages of big data in the file information resources application**

Archivist circle must be kept in the process of archives information is highly sensitive to the development of information technology, new technology trends, timely tracking further research all kinds of information technology is introduced into the possibility of archives work, the necessity and possibility of positive and negative significance [3]. Big data technology and thoughts of archives informatization and the influence of demand is mainly manifested in the following.

**Contribute to the generalization of the concept of file.** In information scarcity and save cost huge paper, selection and save the value of the society and individuals has obvious file is reasonable and necessary. Although a lot of information has a "records", but because of the value of unit information is limited and cannot be saved as "file", such as network access log, track of the mouse on the page, customers in the shopping mall's travel notes, etc.

Scholars generally define the "file" to the end of the event after finishing archive file, and file segmentation of the small, relatively independent information unit. In the information age, the record is archives, once digital record is generated into not deleted files [4]. Widely used in the social practice of tax management system, the household registration management system, such as vehicle traffic management system in real time will "records" into "file", and each record will become archives of long-term preservation value, the unit value of low density record information will serve as an integral whole to the value of archives. In the era of big data, archives concept generalization of exist serious problem, not a cause for concern.

**To break archives information problems.** For a long time, the general layout of archives information is not clear. It is generally believed that "the purpose of the digital archives is two: to protect the archives especially endangered archives and digital copy is made, in order to realize the archives information sharing under the network environment using a digital", and this understanding will archives information computerization of limitations of traditional archives work means and insufficient recognition of the overall value of archives information resources, which is difficult to

dig deep value file [4]. Helping people to understand the objective world, auxiliary decision-making, benefit is the vitality of archives utilization. Big data analytics is not archives staff expertise, but the data analysis provides the basis for large data is of the value of the archives work in the information age.

**Have a profound impact on electronic files work.** In the era of big data, which requires the information "digital" rather than just "digital". Digital refers to the phenomenon into the quantification of index analysis of the process, not just into "0" and "1" in the process of digital formats. Is required for electronic document management with perfect metadata and realize full text search [4]. At present, the electronic document metadata research focused on assure the authenticity of the reality of electronic documents, completeness, readability, security, for the development and utilization of the electronic document may occur in the future. Perfect metadata standard is to provide high quality data for big data analysis and application of the foundation. Based on word frequency and the semantic analysis will be an important means of data analysis, thus, realize the full text input has more important significance. At present, a lot of archives heavy light "digital" and "digital", to image processing of paper files, file format confusion, yuan data collection is not complete, it will become an obstacle to the future of exploitation and utilization of big data. Some scholars one-sided emphasis on electronic document security technology and artificially set many technical protection measures for the electronic file, use will also be added. For the security of electronic documents security should be more emphasis on the improvement of the management system, rather than increasing technology complexity.

### The era of big data school archives information resources service mode of higher education research

Under this background, the school data integration model based on the large data research, data can be unified standard, standard data management, enhance the security of data, improve the efficiency of the use of data resources, make originally scattered, low utilization of education resources are more the value of mining, for schools to provide scientific basis for the decision-making for the development of strategic planning (such as Fig.1). In the light of the data of internal integration, this article from the three dimensions of consolidated pattern design, respectively is the organization dimension, data dimension and technical dimensions.

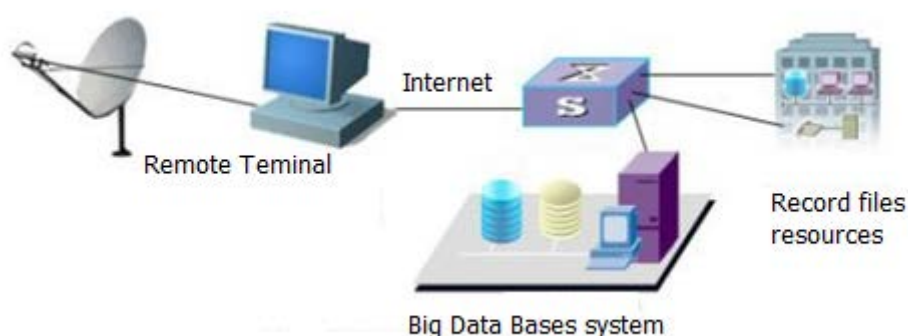


Fig. 1 Information resource sharing model under the big data

**Organization dimension.** Many schools in the early part of the informatization construction invest a lot of money for buying all kinds of application software, system, laying networks, but not to overall planning of informatization, led to a number of departments system is not compatible, cause resource waste, or only pay attention to the information in the form of, don't pay attention to the content, cause a lot of information management module content empty, slow update speed, and is more outstanding is the lack of normative management institution, did not achieve unified planning, unified management, affect the quality and efficiency of the information [5].

Each department in the school can be complementary advantages, realize the integration of the different types of big data resources of high quality. Second internal departments in the school set up

a data management, data integration and management normalization, the agency by the departments in charge of the direct leadership of informatization, and the school leadership team is responsible for data management, and implement the unified deployment of school data integration to coordinate the internal work.

**Data dimension.** To fully tap the potential value of the big data, data integration first clear what data should be included in the scope of transfer, archive, to distinguish the basic data and exclusive data, from the data source to ensure data accuracy and consistency [5]. From the school of business and the existing data, the characteristic that the school in recruitment, employment, education, finance, party and government management and so on various has preserved value formed in the practice of data is the key point of data integration management.

**Technical dimensions.** Data integration is dependent on the technology support, can say the school information construction and application of data is with the continuous development of computer information technology to change [5]. Traditional school of information and data applications use is the unified network platform construction, business platform of the construction of the respective mode, in a few short years, the network infrastructure level of each school got rapid development, core, gigabit Wan Zhao exports [5], gradually establish a data center, the original network bottlenecks, hardware, get a lot of improvement in performance. Contrast is the campus business system and data application in hardware equipment, network bandwidth greatly improved at the same time, not to the school each department's work to effectively improve the quality and efficiency, the role of application system did not achieve the desired effect, and the hardware equipment utilization rate of large investment, high maintenance costs and the contradiction between the low, poor data based growth and geometric data utilization, storage, scattered, poor consistency of contradictions increasingly prominent.

## Summary

This article is based on large data model for data integration, puts forward the school by organizing data, information data, techniques, three dimensions of school integration are studied, the data consolidation work provides a new way of thinking for the school. School level of data integration is not only reflected the school education information level, also reflected the vision of modern education development trend of policy makers, but also measure the abilities of running school and teaching scientific research levels of one of the important standard.

## References

- [1] Zh.L. Liu and J.L. Zhang, Based on the education of information resource integration framework research of cloud, *Intelligence Journal*, 2013, pp.10-16.
- [2] X. M. Wang and J.H. Li, Analyses the concept of information resource education of colleges and universities, *Intelligence Journal*, 2012, pp.54-58.
- [3] L.H. Jiang and L. F. Hai, Big data era scientific data integration research, *Theory and Practice of intelligence*, 2014, pp. 21-26.
- [4] R.J. Tang and Y.X. Zhong, Exploration and practice data center construction, *Information resources of university (natural science edition)*, 2009, 13-18.
- [5] H.B. Li and L.J. Mao, Build research based on data integration of university information standards, *Education information in China*, 2011, pp.10-15.