

Construction and Research of Big Data Platform for Party Building in Colleges and Universities

Yi Han^{1a}, ShuangLiu^{1b*}

¹ School of Marxism, Dalian University of Science and Technology, Dalian, China
hanyi@dlust.edu.cn, Liushuang@dlust.edu.cn

Abstract

The construction of data resources is an important part of the informatization construction of colleges and universities. The construction of a big data management service platform will provide effective support for the construction of big data application analysis scenarios in colleges and universities, including data, management, services, and technology. Taking the DAMA data management service system methodology as the theoretical basis, combined with the characteristics of college party building data, on the basis of the related technologies of the traditional college data sharing platform, the Hadoop framework system, Spark technology, Elasticsearch data engine technology and other big data technologies are deeply integrated. The idea of building a one-stop big data service platform with unified data storage, unified management, unified processing, and unified service provides a reference for the construction of a big data platform for party building in colleges and universities.

Keywords: party building in colleges and universities; big data technology; data platform construction; party building data governance in colleges and universities

1. INTRODUCTION

With the continuous deepening of the informatization construction in colleges and universities, the direction of informatization construction has gradually transitioned from the driving orientation of business system construction to the new goals and needs of data and platform construction content. Utilize the integration of current mature big data technology and traditional database technology, and perform unified storage, integration, calculation analysis, fast query and call of various types of data, fully tap the value of college data, and provide accurate and comprehensive decision-making for the future innovation and development of college party building Analysis will be an important goal of building a big data platform for party building in colleges and universities in the future. This paper integrates traditional relational database and HDFS distributed file system technology, and proposes a methodology for building a big data platform for party building in colleges and universities.

2. ANALYSIS OF THE CURRENT SITUATION OF PARTY BUILDING DATA IN COLLEGES AND UNIVERSITIES

With the gradual improvement and in-depth application of the informatization construction in colleges and universities, the content construction of informatization has gradually enriched. The colleges and universities have completed the construction of campus website homepage and websites of various departments and offices, and developed or purchased some business management information systems or personnel management systems., educational affairs system, scientific research system, etc., but the construction of big data platform for party building in colleges and universities is relatively lagging behind. Therefore, there is an urgent need for a corresponding information system to solve the current problems. In addition, with the requirements of party building in colleges and universities in the new era, data construction and data analysis are also required for party building work and party member development in colleges and universities. These data contain great value, and existing EXCEL statistics cannot process and analysis these data. Therefore, in order to better support the new development of party building in colleges and universities in the future,

colleges and universities need to build a unified party building management platform, which can meet the requirements of unified storage of the existing multi-source heterogeneous data of the school through a set of unified data standards. Governing and standardizing these data, incorporating it into the management system of the data platform, providing data services to the outside world, and providing a solid foundation for future data applications and data mining analysis.

3.CONSTRUCTION IDEAS OF BIG DATA PLATFORM FOR PARTY BUILDING IN COLLEGES AND UNIVERSITIES

Party building in colleges and universities is an independent and complex social ecological system. Because it involves a wide range of businesses, its construction is a huge and complex system engineering. Since the data generated by party building in colleges and universities has the characteristics of various types, complex structures, and wide-ranging business areas, how to meet the current needs of colleges and universities for data-related business construction such as unified data management, data storage, data services, data applications, and data security? It has become an important solution for designers of big data platforms for party building in colleges and universities. Therefore, we should proceed from reality, scientific planning, unified design, and gradual advancement ensure the success of platform construction. The construction of big data platforms in colleges and universities mainly adheres to the following construction ideas:

3.1. Unified planning and design

In view of the complexity and diversity of party building data in colleges and universities, the construction of the party building big data platform in colleges and universities must make a unified and forward-looking scientific plan for the overall structure according to the development needs of the school's future party building work, and form the overall structure of the party building big data platform. The blueprint provides the overall construction basis and goals for the subsequent platform construction.

3.2. Gradually expand the scale

Because the big data platform is a large scale complex system integrated through multiple technologies, it involves a wide range of businesses, and the data is large and complex. Therefore, the construction of the platform cannot be accomplished overnight. It is necessary to gradually expand the scale of the platform according to the overall architecture design and use the method of stages, and finally form a unified university big data platform.

3.3. Focus on infrastructure construction

A good foundation is the key to the construction of high-rise buildings. If the platform wants to have complete functions, a long life cycle and stability, infrastructure construction is essential. Therefore, the content of the construction plan should focus on the selection and construction of basic software and hardware, and the management of basic data resources. and specifications, platform function design, formulation of relevant management specifications and other important factors, and cannot be separated from the basic and practical needs, let alone practical.

4.THE OVERALL ARCHITECTURE DESIGN OF THE BIG DATA PLATFORM FOR PARTY BUILDING IN COLLEGES AND UNIVERSITIES

The overall design of the platform is divided into three parts: technical support platform, data specification platform, and data service platform. Among them, the technical support platform is the underlying software and hardware resource components, the data specification platform is for data managers to perform system configuration, statistical analysis and other operational functional components, and the data service platform provides data operations and services for ordinary users.

4.1. Technical Support Platform

The technical support platform is the basic support for all data services of the big data platform. The platform provides a variety of data query and processing technologies, data storage technologies, and data computing and processing technologies, with scalable and distributed capabilities, including hardware resources, operating systems, data storage, data retrieval, data computing, and data collection. technical components. The technical support platform architecture is shown in Fig. 1.

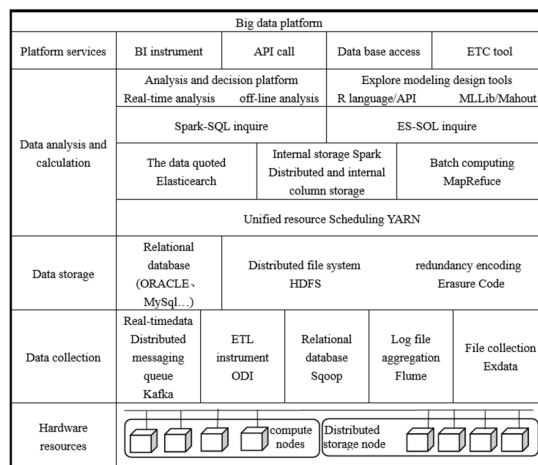


Fig.1 Overall architecture of the technical support platform.

4.2. Data Specification Platform

The data specification platform is a system that provides comprehensive data management support in the big data platform. Users can perform online detection, management and maintenance of data standards, data models, and data interfaces by operating on the front page, and also manage and approve data use applications. The data specification platform architecture is shown in Fig. 2.

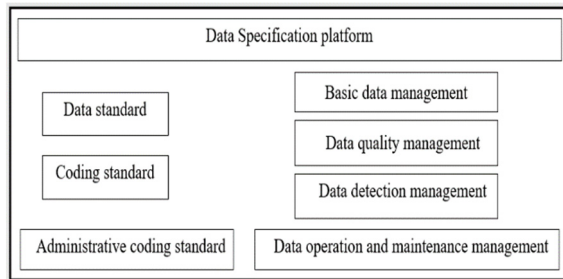


Fig.2 Data specification platform functional architecture

4.3. Data service platform

The functions of the data service platform are divided into the functions of providing users with data authority management, authentication, and data security settings, while data service management provides users with data sharing services, data query services, data analysis services, and data presentation services. The architecture is shown in Fig. 3.

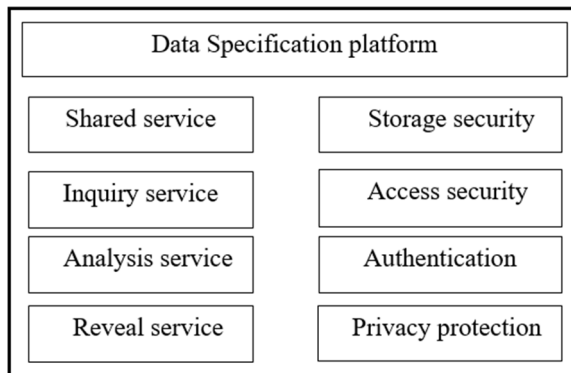


Fig.3 Data specification platform functional architecture.

5. CONTENTS OF BIG DATA PLATFORM CONSTRUCTION IN COLLEGES AND UNIVERSITIES

5.1. Data standard system

The data standard system includes data coding standards, data standards, data code standards and other related standards, and is an indispensable basic support to ensure the standardization of school data, data exchange and sharing, and to give full play to the value of data

resources. Master data coding is a code assigned to information elements in order to facilitate the storage, retrieval and use of information during information processing. Scientific and comprehensive information coding specifications are an important part of the informatization standard system, the premise of information sharing and sharing, and the indispensable basic support for ensuring data sharing and exchange and giving full play to the value of data resources. According to the main data types of the core business of colleges and universities, it can be divided into 12 categories of coding rules, such as organization code, staff number code, student number code, major code, class code, campus code, and project code. The data standard is a data standard covering the school profile, students (undergraduates, graduate students), personnel, scientific research, finance, assets, office, foreign affairs, archives, books and other information of each branch of the school. The basis of school design data standards should be based on the "JYT1002_Educational Management Basic Information", "JYT1003_Educational Administrative Management Information", "JYT1006_Higher School Management Information" issued by the Ministry of Education in 2012, combined with the school's specific party building work. As shown in Fig. 4.

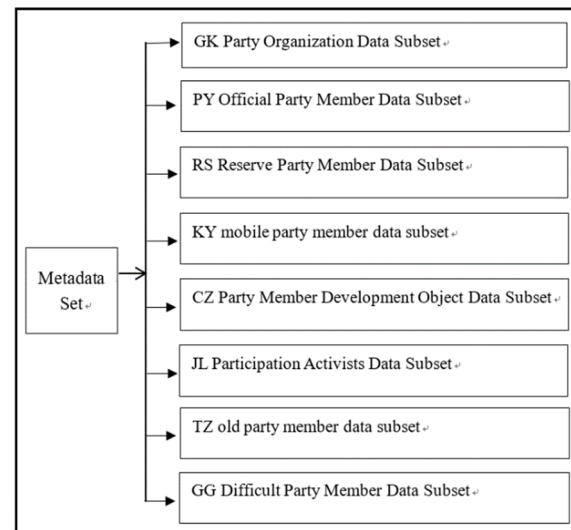


Fig.4 Metadata Standard Classification

The standard code is based on the national standard code and the education industry standard code, combined with the specific situation of the school, and according to the classification specification of the metadata set, the code standard compilation work is carried out. Among them, 51 kinds of national general standard codes are cited, 148 kinds of standard codes of the Ministry of Education are cited, and 31 kinds of school code standards are formulated.

5.1.1. National standard

The data code standards for colleges and universities include 51 types of public data code standards such as

gender, certificate type, blood type, and political affiliation, all of which use the national unified data code standards, namely GB/T2260, GB/T2261.1-GB/T2261.7, etc. For details, please refer to JYT0001_Educational Management Fundamentals Code.

5.1.2. Education industry standards

The codes of the college education industry in colleges and universities, including degree category, tutor category, faculty category, etc., all adopt the education industry consent data code. For details, please refer to JYT0001_Education Management Basic Code.

5.1.3. School standards

College standards are based on the expansion and modification of national standards. The reason for using school standards is that the code types provided by the state cannot cover the needs of schools, or there are no related code types such as college employment contract types, card types, and most codes for logistics. Such school standard codes need to be formulated according to the specific circumstances of the school.

5.2. Data Resource Center

The data resource center is the storage unit of the university party building big data platform that centrally and uniformly stores various party building data of the whole university. It is mainly used to store structured data generated by traditional business systems and offline unstructured document data (mainly Excel documents, pictures, Word documents, etc.), log data generated by network devices or systems, geographic information data, Internet of Things data, etc. It is the basic support for the core of the big data application service of the whole school.

5.2.1. Data Storage

Technically, there are many kinds of complex data structures, so the data resource center is a complex storage module integrated with various types of data storage technologies. Traditional business system data should be stored in traditional relational databases, and other semi-structured and unstructured data should be stored in HDFS. Logically, the data resource center needs to collect and store data for storage management. Therefore, no matter what type of data, it should be planned into different logical storage areas according to the data storage strategy to facilitate subsequent management and invocation.

5.2.2. Specification of Data Cleaning and Transformation

Unified data storage is the basis for ensuring the utilization of data resources, and data quality is an

important guarantee for whether the big data platform can play the value of data. Heterogeneous data cleaning, data integration, and multi-type data conversion are the prerequisites for ensuring data quality. Regular inspections of data such as data timeliness, data integrity, data accuracy, and data standardization are the basis for data quality assurance. In order to improve the quality of data, it is necessary to formulate a set of data cleaning and conversion specifications based on the relevant content of the data standard system and the actual situation of party building data in colleges and universities.

5.3. Data Service Center

The early data exchange platform in colleges and universities is only a service platform that provides data exchange and sharing for various units in the school. Technically, most of them are pushed on demand by the information department through interface tools such as ETL. The data service center is a comprehensive upgrade of the original data exchange platform. In addition to retaining the original data exchange function, it also needs to expand its functions in terms of service scope, supporting technology, and service functions. Among them, the scope of services has expanded from the original data sharing only for business departments to personal data application services such as students or teachers through online application. Technically, based on a single ETL technology, new functions such as API interface calls and SQL view calls are added, and all data interface calls are unified in a single system interface, so that data interface users can only operate the foreground application layer without the need for Understand how the background data is generated. In addition to the data sharing function, the service function also provides online data calculation function, data BI display function and data query and retrieval function. It makes the data usage more active and maximizes the value of the data.

5.4. Data Security System

5.4.1. Authentication and Auditing

Authentication. User identity authentication includes authentication of data users through technical means such as passwords, fingerprints, and security keys, so as to prevent others from using their identities to access relevant data.

Access control. Access control is performed on the access IP address and access window, generally using IP address restriction access or bastion machine access.

Security audit. All operations in data access are recorded to realize prevention, in-event early warning, and post-event audit.

Privacy management. Data privacy management is the realization of security management for the sensitive

data of general data individuals or a certain group, including requirements such as data release prevention and user group division.

Data release anti-leakage. Pay attention to security when publishing data, and do not leak data during or after publishing. Specifically, you can obtain the consent of the data owner before the data is published, and encrypt sensitive and unimportant data.

Division of user groups. Divide users into multiple groups according to the level of data sensitivity and the permissions the data has, and there must be no mixing of data across groups.

5.4.2. Data storage security

Data storage security is to achieve data security storage through data storage security technology, generally realized by data encryption, usually data encryption can use plaintext encryption or algorithm encryption.

Data access security is to ensure secure data access by ensuring terminal security. The server terminal used for data access should undergo security scanning inspection, and conduct security audit through the auditing system to prevent data leakage due to malware tools such as Trojan horses or viruses being implanted in the terminal during data access.

6.PLATFORM CONSTRUCTION PRINCIPLES

The architecture adjustment design of big data platform should follow the following basic principles.

6.1. The principle of practicality

The big data platform must first adhere to the principle of practicality, consider advanced and forward-looking on the basis of practicality, and effectively meet the actual work needs of the school's big data application business.

6.2. Security principle

The big data platform support system follows the relevant requirements of the school information security system, and adopts relevant security mechanisms and technical means to ensure the application security, data security, server security, network security and physical security of the system.

6.3. The principle of reliability

The big data data platform should meet the requirements of 7 d × 24 h reliable operation of business applications. The software and hardware resources of the key links of the system are designed with high availability solutions to ensure high reliability of system operation.

6.4. The principle of scalability

The party building big data platform should adopt a flexible design, have good scalability, and have flexible configuration capabilities for party building work processing.

7.CONCLUSIONS

In summary, the application of the university party building big data platform has the following core contents:

- 1) The university party building big data platform must be a centralized storage platform for the entire university party building data information storage platform, providing a solid basic data support for the comprehensive application and analysis of university data.
- 2) The relevant data construction of the platform must have a set of data standard system that conforms to the actual situation of the school as a unified guiding standard.
- 3) The technical architecture of the big data platform for party building in colleges and universities should have both the ability to process traditional shared data and the ability to process unstructured data, and at the same time, it should have a set of perfect and practical data management service specifications to support its management, operation and maintenance.
- 4) During the construction and operation of the university party building big data platform, security protection should be carried out in several aspects such as user authentication, data storage, and data access. The idea of building a big data platform for party building in colleges and universities proposed in this paper not only inherits the advantages of traditional database technology to solve the data sharing business in colleges and universities, but also solves the problem of unstructured data storage and processing. The data application platform provides practical and feasible construction ideas.

ACKNOWLEDGEMENTS

This work was supported by the Scientific Research Project of Education Working Committee of the CPC Liaoning Provincial Committee and Liaoning Provincial Department of Education(2021GXDJ-YB089); The 2021 annual Scientific Research Funding Project of the Liaoning Provincial Department of Education (LJKR0692); Research project of Dalian Academy of Social Sciences(2021dlsky188).

REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (references)

- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

