International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 2, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n2.22
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Research on University Data Governance Scheme Based on Data Platform

Minfu Xu [1, *], Zhixiang Zhuang [2], Junjun Wu [2]

[1] Network and Information Office, North China Electric Power University, Baoding, China
[2] Computer Science Department, North China Electric Power University, Baoding, China
*Corresponding Author: Minfu Xu

ABSTRACT

A data governance solution for universities based on a data center is proposed to address various data management issues encountered in the construction of smart campuses. The solution consists of three parts: a data base, a data resource system, and a data center. With data assetization as the core and data serviceization as the goal, a complete set of mechanisms including data collection, standardization, cleaning and conversion, supporting management, and interface openness are adopted to cover the full lifecycle management of data, explore the value behind data, achieve synchronous improvement of data quality and data service quality, and promote the development of information technology and smart campus construction in universities.

KEYWORDS

Data governance; Data center; Information construction; Data

1. INTRODUCTION

The development goal of informationization in universities has always been to build a “digital campus” for business system informationization and data resource collection. After decades of development, building a “smart campus” has gradually become the goal of university development [1]. The “digital campus” is the foundation of the “smart campus” and an inevitable manifestation of the development of the “digital campus” to a certain stage [2]. The "smart campus" integrates and analyzes a large amount of application data accumulated by various business systems and software and hardware devices of universities in the "digital campus" stage through information technology, and uses it for deep application, intelligent process, personalized service, and overall improvement. By innovatively integrating data resources with information technology, it can continuously improve the level of campus information construction and application. The guiding documents such as the Action Plan for Education Informatization 2.0, the Notice of the Ministry of Education on Strengthening the Informationization of Education Management in the New Era, and the Norms for the Construction of Digital Campuses in Higher Education Institutions have pointed out the importance of comprehensively promoting the construction and application of smart campuses in schools [3]. Building smart campuses has become an inevitable trend in the development of university construction in the new era.

The rapid development of big data technology has enabled it to be widely applied in various fields. By conducting big data analysis and mining on the massive data accumulated during the construction of the "digital campus" in universities, various applications can be developed for campus governance, education and teaching, logistics management, and other aspects [4]. A new model of university

governance can be constructed to enhance the digital level of education management and assist in the construction of a smart campus in the new era.

The data center is a platform and mechanism established to continuously improve data quality and data service capabilities. Through the data center, data scattered on various information islands can be integrated to form stronger data service capabilities, providing analysis and decision support for managers. Therefore, the development of informationization construction in universities in the new era should be based on the principles of "unified standards, top-down linkage, and resource sharing", establish a data center, do a good job in data governance, break down data barriers between business systems, and build a directory and traceability map of data and information resources. Promote the orderly sharing of educational data, expand the scope of data sharing, improve data sharing efficiency, better support the service application of business departments, gradually promote the process of business dataization, data assetization, and asset value realization, and use data governance to promote the optimization of higher education governance capabilities.

2. CURRENT SITUATION OF DATA MANAGEMENT IN UNIVERSITIES

The data standards are not unified, and the construction of various management information systems in universities is different. The industry has not formed a development background that follows a unified data standard specification. Each vendor adopts its own data definition, code set, measurement unit, expression format, coding method, etc. The lack of uniformity in standards and specifications between data leads to a large amount of data being in a state of isolation and isolation, making it difficult to achieve correlation aggregation and big data analysis [5]. The data source is not clear, and in the actual management scenarios of universities, there are often multiple sources of data. When the data in one system changes, there is no corresponding mechanism for synchronizing the corresponding data in other systems. After a period of operation, the data of each system begins to become different, making it difficult to determine which data is completely correct and complete. In this case, it is impossible to build a high-quality dataset that supports the operation of a smart campus. Due to various reasons such as functional defects, lack of management, improper operation, etc., there are often various problems with data quality during the operation of various systems. For example, missing data items, poorly structured data, invalid enumeration items, incorrect expression formats, numerical errors, etc [6]. Difficulty in data coordination and sharing. Each department often needs to use data generated by other departments in their own management process, but in actual data coordination work, they often face the "black box dilemma" [7], which makes data acquisition very difficult. There is security risks associated with data sharing, and the existing data supply methods mainly rely on manual querying, summarizing, and organizing at the bottom of the database, or sharing and using data through offline tables. This means that online data usage cannot be traced, and data security cannot be guaranteed. At the same time, improper use of data cannot be detected and prevented in a timely manner, posing a significant threat to data security management.

3. THE CONSTRUCTION GOALS OF DATA GOVERNANCE IN TWO UNIVERSITIES

During the operation of information management systems, a large amount of data is generated, which can be used to serve business management fields such as education and teaching, student management, and scientific research. Therefore, driven by data, building a data resource system and management system, strengthening the construction of information services, supporting innovative business management, personalized talent cultivation, scientific research and other service management methods in schools with information services, and further improving the governance system and governance capacity construction of schools are the current construction goals of data governance in universities [8].

3.1. Construction of a Global Data Center

The construction of data governance should take people, finance, and materials as the starting points, carefully sort out the information management systems of universities such as personnel system, student engineering system, graduate system, OA office, academic system, scientific research system, financial system, etc. Through comprehensive data collection, centralization, and governance, the expression of data should be standardized, the content accurate, and the dimensions complete, forming a standardized, classified, and reliable data warehouse and data mart.

3.2. Establish a Data Lifecycle Management System

The full lifecycle management of data resources includes various aspects such as data collection, data warehouse management, data interface opening, data resource publishing, and service status presentation. Integrating and managing the collected data not only enriches the storage of data resources, but also provides the latest support for new service applications, ensuring the accuracy of service applications. The warehouse management of data involves a series of operations and management of stored data resources, timely updating of data status, ensuring the authenticity and real-time performance of data, which is crucial for subsequent data applications. Only by collecting and managing data resources well can we lay a solid foundation for subsequent data interface opening, resource publishing, and other processes. When opening data interfaces, secure, advanced, and convenient technological means should be used, including programming interfaces, database interfaces, message interfaces, file interfaces, and other online and offline methods, to provide services for various data application scenarios.

3.3. Personal Data Visible, Manageable, and Usable

Quickly assemble all relevant personal data of teachers and students, including basic information, teaching, research, assets, daily life, and other data resources, into a personal data center to achieve data and personal communication; Provide data management, error correction, and supplementary recording functions based on personal data centers to address closed-loop data quality issues; Based on personal data, a business filling form is constructed to solve the problem of difficult data acquisition and repeated filling in for teachers and students when filling in data. This reduces the workload of teachers and students in filling in data, while achieving visibility, management, and utilization of personal data.

3.4. Provide Data Support for Comprehensive Digitization

In the daily operation, management, and decision-making process of schools, almost every activity is closely related to data. After the completion of the global data center, it can better connect with other departments' data resource sharing and exchange with the same standard, and provide data services for various scenarios such as the teacher-student service hall, mobile campus, information portal, microservices, and data filling. Through accurate and detailed data, it ensures the smooth operation of various businesses, reduces the energy and cost of data coordination, and helps the leadership to timely and accurately grasp the various situations of the school through data analysis, providing data support for management and decision-making processes.

4. UNIVERSITY DATA GOVERNANCE PLANS

The data governance solution for the data center is a specific and practical solution for university data centers. Its construction goal is to establish school level data standards, complete the confirmation of "one data source", build a school level data resource platform with rich content, complete dimensions, clear hierarchy, and reliable quality, achieve full lifecycle management of data resources, and

ultimately provide high-quality data services for various information services in the school, promoting the rapid implementation of various information services driven by data. From a process perspective, the entire process of data governance is shown in Figure 1. Firstly, it is necessary to conduct a survey of data resources, understand the overall situation of school data resources, determine the sources, quantities, types, and access methods of data, and then carry out data collection. During the collection process, it is necessary to extensively and comprehensively collect data from various application systems, electronic documents, paper files, and other carriers in the school; Next, we will build a data resource system, establish school level data governance standards, clarify data rules, clean and transform data, ensure data quality, and do a good job in data retention and classification; Finally, configure data management tools to meet the different data operations of various personnel in daily work, ensure data privacy and security, open various ports, and interface with various application systems on campus.

Figure 1. Data Governance Process Diagram

4.1. Data Base

The data base specifies the storage methods for different types of formatted data and is the most fundamental part of the entire data governance solution,

It mainly consists of two types of infrastructure, relational database and big data architecture. Relational databases are used to store structured data, corresponding to data from various business systems and spreadsheets, as well as result data tables generated through statistical calculations; Big data architecture is mainly used to store large amounts of unstructured data, corresponding to network log data (including raw and parsed forms), as well as some large amounts of structured data, such as one card consumption flow tables.

4.2. Data Resource System

The data resource system is the result of data resources generated through data governance services, and is the most core and important part of the entire data platform. Data resources are divided into three components: data lake, standardized data warehouse, and data mart. After data governance in the data resource system, data standards have been unified, data sources have been clarified, and business silos have been avoided; The data quality has been improved, reducing the occurrence of missing, erroneous, and omitted data; Solved the problems of difficult data coordination and insecure data sharing. The data resource system will be integrated with various smart campus applications to support the operation of various applications and analyses. The flow of data resources throughout the school is shown in Figure 2.

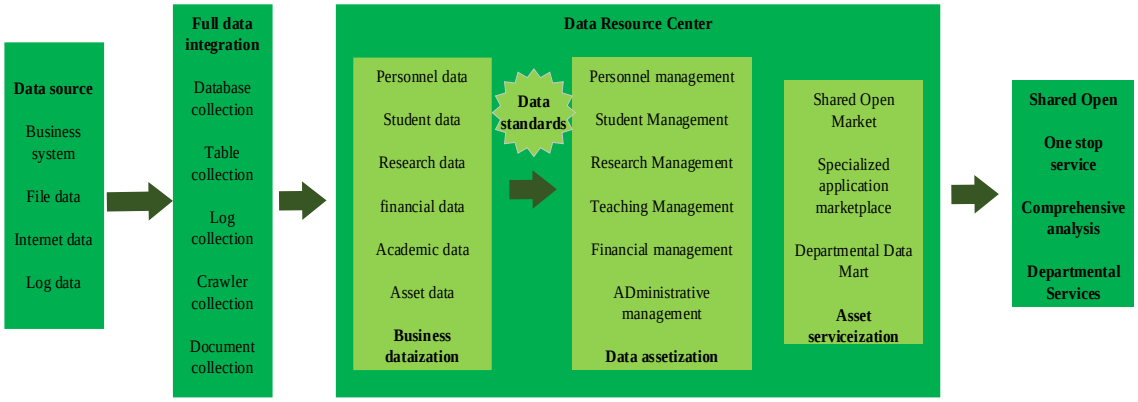

Figure 2. Schematic Diagram of Data Resource Flow

4.3. Data Center

The data center is the core part and important tool of data governance solutions. It is not only a technical platform, but also the external implementation of data service capabilities. Through data services, it directly promotes the operation of various business applications in schools. So, the data center cannot only target professional and technical personnel, but also requires the participation and collaboration of management and business personnel from functional departments of the school.

The data center mainly completes data collection, including the collection of various forms of data, online filling of missing data, and collection and management of metadata; Data processing, converting and cleaning data; Data assetization, systematic management of data structure, content, quality, and interfaces according to data standards; Data application enables the application, authorization, and integration of data resources with software, as well as the ability to view the usage of monitoring data, enabling autonomous monitoring, operation, service, and full lifecycle management of university data resources. The functional architecture of the data center is shown in Figure 3.

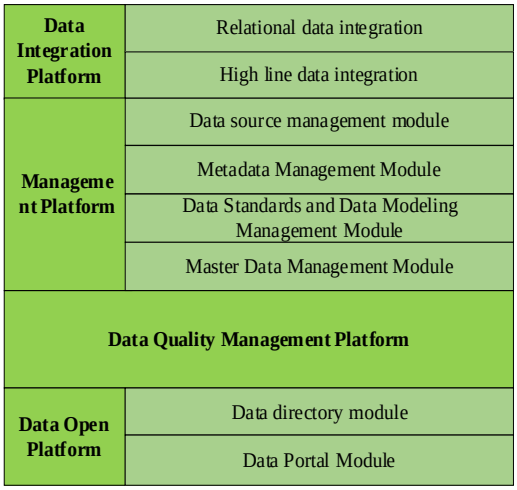

Figure 3. Functional Architecture Diagram of Data Center

5. IMPLEMENTATION EFFECT OF DATA GOVERNANCE

The comprehensive data warehouse formed by data governance will provide complete and accurate reference data for all departments, leaders at all levels, and teachers and students in the school. Fine management measures can be formulated based on the results of comprehensive data analysis. A full data warehouse can also fully connect the remaining data silos, achieve the full flow and utilization of data, continuously correct data errors, generate more new data, and continuously enrich and improve the content of the data warehouse, achieving a virtuous cycle and sustainable development. In future business management software development work, a full data warehouse can provide standardized basic data and achieve exchange and sharing with other system data, thus significantly reducing the cost of system integration and reducing the fees charged by developers for preparing data.

At the same time, the data center has value for service recipients at all levels, such as providing data resource support for various information services on campus, promoting the interconnection and linkage of various information services on campus, and accelerating the intelligent process of campus management, office, and services; Provide available data for management optimization and decision-making; Improve the user experience of one-stop services; Provide data for reports and reduce duplicate reporting. The informationization capability of universities, the service level of

informationization management departments, and the informationization service experience of teachers and students have also been significantly improved.

6. CONCLUSION

Data is a new type of production factor following land, labor, capital, and technology, playing a crucial role in the construction of smart campuses in today's era. The purpose of data governance is to explore and enhance the value of data resources, which can drive the development of various business systems and application services in universities with data, and further promote the informationization construction of universities in the new era. The data governance solution proposed in this article, consisting of a data base, a data resource system, and a data governance and service platform, covers the entire lifecycle management from data collection to external data release and application. It can effectively improve the quality of data, standardize, asset-backed, and service-oriented data, and explore the various data values accumulated by universities for a long time. However, it should also be recognized that data governance cannot be achieved solely through a comprehensive plan. What is more important is the construction of supporting school management systems, follow-up planning of business processes, and training of personnel literacy at all levels. Only in this way can the data governance plan be better implemented and the maximum value behind the data be realized.

REFERENCES

[1] Wu Ruifang. “Research and Exploration on the Construction of Smart Campus in Universities under the Background of Big Data”, China Information Technology, Vol. 9, pp. 94-95+101, 2023. https://doi.org/10.3969/j.issn.1672-5158.2023.09.044

[2] Chen Yan. “Research on Data Governance of University Informatization”, Network Security Technology and Application, Vol. 3, pp. 73-74, 2023. https://doi.org/10.3969/j.issn.1009-6833.2023.03.034

[3] Tu Jiaqi, Wang Dongmei, Gao Huanjiang, et al. “Research on Big Data Service System in Universities under the Background of Smart Campus”, Modern Electronic Technology, Vol. 20, pp. 76-80, 2023. https://doi.org/10.16652/j.issn.1004-373x.2023.20.014

[4] He Jianhu. “Design and Implementation of Data Governance Architecture for Smart Campus in Universities”, Information Technology and Informatization, Vol. 6, pp. 157-160, 2022. https://doi.org/10.3969/j.issn.1672-9528.2022.06.040

[5] Li Chaofan, Sun Xiaoxu. “Research on Innovative Practices of University Governance under the Background of Big Data”, Inner Mongolia Science and Technology and Economy, Vol. 2, pp. 41-43, 2024. https://doi.org/10.3969/j.issn.1007-6921.2024.2.nmgkjyjj202402009

[6] Zhang Tian, Zou Henghua, Shao Weihui, et al. “Exploration of University Data Governance Based on Big Data Data Warehouse: A Case Study of Tongji University”, Modern Educational Technology, Vol. 2, pp. 118-125, 2023. https://doi.org/10.3969/j.issn.1009-8097.2023.09.012

[7] He Yuange, Mu Rong, Zhao Anxin. “Research and Practice of University Data Service Framework Based on Big Data Governance: A Case Study of Xi'an University of Science and Technology”, Modern Information Technology, Vol. 23, pp. 25-27, 2022. https://doi.org/10.19850/j.cnki. 2096-4706.2022.23.007

[8] Wei Jianxing, Liu Yuanzhi, Luo Chao, et al. “Research on University Data Governance System Based on Data Platform”, Information Technology and Informatization, Vol. 6, pp. 98-101, https://doi.org/2022.10.3969/j.issn.1672-9528.2022.06.024