

Research on the Application of Data Governance for College Teachers

Xiaofei Li

Jilin University of Architecture and Technology, Changchun 130012, China

Abstract: In recent years, many universities have embarked on the construction of the “One Form” project through one-stop service halls to address the challenge of teachers’ difficulties in filling out forms. However, in practical applications, data quality issues are prominent due to factors such as non-uniform data standards and untimely data maintenance, severely impacting teachers’ data application experience. This paper, through analyzing the current situation of teachers’ data quality, proposes a solution for teachers’ data governance based on the one-stop service platform. By establishing data standards, building a standard data warehouse, and instituting a data management guarantee system, this paper aims to continuously improve teachers’ data quality and enhance their data service experience.

Keywords: Data governance; Data standards; Data quality

Introduction

Data serves as the foundation for modernizing educational governance. Making data standardized, accurate, unified, shareable, and interoperable through data governance is a prerequisite for data to serve university management and development. Some scholars argue that data governance is crucial for improving educational quality, decision-making scientificity, and management efficiency. Effective application of data governance can make university management more intelligent and agile. As the mainstay of teaching and research in universities, teachers are the primary producers of data. As the subject of data, teachers should not be absent from data governance. Their participation and feedback can effectively drive the data governance process and ensure its effectiveness. In recent years, many universities have successively established university-level data centers and actively practiced system-level data sharing, laying a technical foundation for data sharing and exchange. On this basis, how to promote data governance practices centered on university teachers is an issue that urgently needs to be explored and resolved.

1. Problem Analysis and Governance Strategies

To address the pain point of teachers having to repeatedly fill out forms in their daily work, an investigation and analysis of teacher data has revealed that the quality issues primarily manifest in the following aspects: Firstly, there is a lack of uniform data standards. Over the past decade of information technology construction, the school has successively launched multiple application systems for teacher service management at different times. During the actual application process, due to differences in business jurisdictions, the definitions and understanding of the same data items vary, leading to deviations. Secondly, the data sources are unclear, with the phenomenon of “multiple entries” existing. The same data item may have multiple sources, and inconsistencies exist among these sources, resulting in no definitive authoritative source for the data. Thirdly, data is incomplete, with missing data items, causing teachers to need to repeatedly fill in information. Fourthly, data maintenance and updates are not timely, leading to outdated and unusable data. Fifthly, data entry is not standardized, resulting in invalid data, such as incorrect number of digits for ID numbers or phone numbers, and incorrect date of birth formats. Sixthly, the school has not established a standard data management mechanism, lacking effective data change and review processes.

This study believes that the practice of teacher data governance needs to be considered from three dimensions: data, personnel, and system. Each dimension encompasses several key elements, which collectively form the practical path to promoting teacher data governance in colleges and universities. The specific analysis is as follows:

1.1 Data Dimension

Data Standards Unification: The standardization of data ensures consistency in the collection, processing, exchange, and transmission of field information across various business systems, facilitating information and data resource sharing. **Data Quality Monitoring and Evaluation:** Establishing a comprehensive data quality evaluation system covering dimensions such as completeness, timeliness, and accuracy to ensure

the overall quality of data and provide precise decision-making analytics. **Data Warehouse Construction:** Building a standardized data warehouse based on the school's public data center to achieve centralized data management and integration, enabling the data center to serve as the sole outlet for basic data across the school.

1.2 Personnel Dimension

Clear Responsibility Allocation: Defining clear responsibilities for data generation, management, and usage based on the principle of "who generates the data, who is responsible for managing it." **Teacher Engagement and Awareness:** Encouraging teachers to actively maintain the timeliness, standardization, and completeness of their personal data through mechanisms such as national data reporting and teacher performance evaluations. **Training and Support:** Providing training and support to relevant personnel to ensure they are equipped with the necessary skills and knowledge to effectively participate in data governance activities.

1.3 System Dimension

Governance Framework and Policies: Establishing a comprehensive data governance framework and policies that outline the overall goals, strategies, and procedures for data management. **Data Management Platform:** Developing a data application service management platform to streamline data access and usage, improving user convenience while ensuring data security and traceability. **Integration and Collaboration:** Fostering integration and collaboration across different departments and systems to ensure seamless data flow and consistency across various applications.

These key elements, working in concert, constitute a comprehensive approach to promoting teacher data governance in colleges and universities, enhancing data quality, and ultimately improving the overall efficiency and effectiveness of teaching, research, and management activities.

2. Implementation Pathway

2.1 Data Source and Process Sorting

The sorting of data sources involves an inventory of the existing data assets within the school. Data arises from the process of business management or information service flow. Through an investigation of the school's business processes and the current status of data, the flow of data between systems is mapped out, and the movement, transfer, processing, and storage of data are analyzed comprehensively. The purpose of data flow analysis is to identify and address issues in data circulation, including data completeness, consistency, accuracy, standardization, timeliness, and processing procedures. By sorting through the teacher service processes, the school determines the data items involved in each service process and process node, organizes them by data classification. Through data service classification, teacher-related data can be divided into categories such as personnel basic information, teaching tasks, research awards, and training. By sorting through the relevant data of teacher service processes, a clear picture can be obtained of the source of each type of data, the method of data provision, which data is problematic, and what data needs to be collected additionally, laying a solid foundation for the subsequent standardization of school data. In the specific implementation phase, it can be divided into two categories for targeted handling: First, when teacher data sources are clear and supported by a system, it is necessary to find the corresponding data table or data interface in the system, establish a data relationship archive, and clarify the data content, rules, and synchronization mechanisms. Second, when teacher data sources are unclear and lack system support, it is necessary to plan and design the data collection items and methods based on the school's data application planning, taking into account future business development needs and actual application scenarios.

2.2 Establishment of Data Standards System

The establishment of a data standards system is the core content of data governance. Standardized data specifications have become a basic guarantee for the overall planning and construction of school informatization. The establishment of data standards ensures that the field information of various business systems in the school follows unified standards and specifications during the processes of collection, processing, exchange, and transmission, thereby facilitating the sharing of information and data resources. This lays a foundation for the establishment of data sharing platforms and standard data warehouses based on standard specifications. The focus of data standards development is to formulate a data standards and specifications system. The formulation of data standards and specifications should firstly follow the national Ministry of Education data standards. For fields without national or industry standards, reference should be made to provincial code standards or information data standards that meet the actual needs of the school. The establishment of data standards should fully consider the school's future development and educational and teaching reforms, with a certain degree of scalability, such as the gradual expansion of coding categories and data content along with the school's business development changes and the increase in basic data construction content. Teacher data standards are an important part of the school's data standards system, which can be classified and defined according to actual business needs and the school's future development requirements, encompassing teacher basic information, teaching, and research information, while meeting

both the Ministry of Education's data reporting standards and norms and the school's internal data sharing norms.

2.3 Construction of a Standard Data Warehouse

The construction of a standard data warehouse is the establishment of a standardized dataset built upon the school's public data center. If the school has not yet established a public data center, it should first extract the basic data from major business systems to construct a centralized public data center for the entire school. This allows for centralized data management and integration at the data layer, making the data center the sole source of foundational data across the entire school. Based on this data center, data cleaning is performed in accordance with data standards to create datasets for different application scenarios oriented towards serving teachers. For data that is not supported by business systems, it is necessary to establish the data table structures in advance based on the data standards, in preparation for subsequent data collection. Simultaneously, data history and production databases are built according to the data lifecycle to provide data support for data application and analysis, as well as shared data access services for relevant application systems.

2.4 Construction of a Data Application and Service Management Platform

The construction of a data application and service management platform involves establishing a centralized platform that facilitates user access to data while ensuring data security. This platform enables users to request data access through a streamlined process, where requests are approved by relevant personnel and system-allocated permissions are granted directly. Additionally, the platform is designed to provide a one-stop service flow for teachers based on their data application scenarios, enhancing user convenience while maintaining data security and traceability. This approach not only improves the ease of data access for users but also ensures the security and reliability of data usage.

3. Effectiveness of Data Governance

3.1 Improvement of Data Assets

In terms of data construction, the platform revolves around teachers' teaching, research, and other activities to clarify data assets. All types of data are stored in a standardized format in the database, and maintained and managed according to standards to ensure data normativity. Among the 74 data items constructed, 16 data items are provided by business departments through sharing, while the remaining 58 data items are sourceless. These sourceless data are not managed by any business system and are not stored in a standardized and informatized manner. When needed, they can only be repeatedly filled out, and cannot be shared or reused. To address this issue, the platform has uniformly collected over 80,000 pieces of data through college entry and teacher submission. The platform collects data that are currently not managed by any business system within the school and are not shared to the school-level data center, and manages them comprehensively. This further enhances the comprehensiveness of the data. While improving the school's data assets, it also enables data sharing with other business departments, solving problems such as repeated data collection and reporting.

3.2 Governance of Challenging Data

The collection and governance of research achievements such as papers and patents pose significant challenges in data governance. This type of data is often generated outside the school and not processed through college business departments, making it difficult to collect. Even if data is collected from external institutions, it is challenging to accurately associate the achievements with internal school units and their authors based solely on the raw data collected. Furthermore, the evaluation information of these achievements is dynamic and cannot be directly obtained.

To address the issue of achievement data governance, the "One-Stop Platform" has collaborated with the university library. Through shared access with the library's institutional knowledge base, each achievement is verified by the college and the teacher themselves, and then corrected by the library. The current data accurately reflects the basic information of the achievements, publication details, authors and their roles, affiliated colleges, partitions, high citations, and other information, providing a data foundation for subsequent analysis and applications.

3.3 Enhancement of Data Quality

In terms of data management, teachers conduct a thorough check of their various types of data through the data verification and submission service, identifying and filling in missing information, and providing feedback or reporting errors for any data that is in dispute. In terms of the management mechanism, the platform adheres to the principle of "who enters, who modifies; who generates, who is responsible." It insists on data review, source modification, and multi-point sharing. Through teacher verification, college review, and source modification, the accuracy, authority, and consistency of the data are ensured.

4. Conclusions

Data governance has always been a challenging issue faced by universities in the information age. Issues such as incomplete data, decentralized storage, inconsistent standards, and low quality have made it difficult to effectively utilize data and leverage its value to serve teachers and universities. This article conducts a thorough research and exploration into teacher data governance and finds that the application of the

platform has achieved initial results. The platform positions teachers as a crucial link in data governance, ultimately returning data to individual teachers, and their deep involvement significantly improves data accuracy and enhances data vitality. For colleges, as data managers, they oversee and maintain the data lifecycle, achieving data accumulation, optimizing management processes, and improving management efficiency. Furthermore, the platform fosters a virtuous cycle of data governance by innovating management mechanisms, providing technical support, collecting and governing data, and then serving teachers and university units through data sharing or application.

References

- [1] DAMA International. The DAMA guide to the data management body of knowledge[M]. Bradley Beach, NJ: Technics Publications, 2019:7-12.
- [2] Education Research & Data Center. Data governance[OL]. <<https://erdc.wa.gov>>
- [3] Weber K, Otto B, Osterle H. One size does not fit all—A contingency approach to data governance[J]. Journal of Data and Information Quality, 2019, (1):1-27.
- [4] Malik P. Governing big data: Principles and practices[J]. IBM Journal of Research and Development, 2020, (3/4):1-13.
- [5] Khatri V, Brown C V. Designing data governance[J]. Communications of the ACM, 2010, (1):148-152.
- [6] Seiner R S. Non-Invasive framework for data governance implementation-Part 1: Levels and components[OL]. <<https://tdan.com/non-invasive-framework-for-data-governance-implementation-part-1-levels-and-components/19986>>
- [7] Thomas G. DGI data governance framework[OL]. <<https://datagovernance.com/the-dgi-data-governance-framework/>>

This work was supported by the Jilin Province Education Science Planning Project, Project name: Research on Data of Office Governance for College Teachers, Grant number: ZD19137.