ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

Exploration of Data Governance Achievements and Application in Higher Vocational Colleges from the Perspective of Big Data

Wang Chuan

Sichuan Vocational College of Chemical Industry, Luzhou, Sichuan, China

Abstract: *In the era of big data, higher vocational colleges are faced with the challenge of generating and managing massive data. In order to deal with these data effectively and realize data governance, higher vocational colleges actively explore various results and application strategies. In this macro background, this study takes the data governance direction of higher vocational colleges from the perspective of big data as the entry point, and analyzes its governance results and application cases through literature research, practical research, cluster analysis and other forms, aiming to further deepen the research on data governance of higher vocational colleges and explore more effective methods and technologies. To provide higher vocational colleges with better data support and decision-making basis.*

Keywords: *Big data; Higher vocational education; Data governance; Cluster analysis*

1. Introduction

Data governance in higher vocational colleges from the perspective of big data is an important issue in the current education field. With the rapid development and application of information technology, higher vocational colleges are faced with a large number of data generation and management challenges. Therefore, it is very important for higher vocational colleges to establish a scientific and effective data governance system. Based on this, this study continues to deepen the exploration of data governance achievements and applications in higher vocational colleges from the perspective of big data, aiming to help higher vocational colleges establish a scientific and effective data management system, make full use of big data technology for analysis and application, and provide strong support for education reform and innovation. This is also an important starting point of this study.

2. Value motivation of data governance in higher vocational colleges

As an educational institution, higher vocational colleges produce a large amount of student information, teacher information, curriculum information and other types of data every day. These data are the basic resources necessary for the normal operation and management of higher vocational colleges. By managing these data, its quality and integrity can be ensured, and data reliability and accuracy can be improved. The so-called data governance in higher vocational colleges refers to a series of measures and methods to manage and process data in higher vocational colleges. With the rapid development of information technology, higher vocational colleges are faced with an increasing amount and complexity of data, so data governance has become an important task.

At the same time, it should also be noted that the decision-making process of higher vocational colleges needs to rely on a large amount of data support. Through the data analysis and mining of student enrollment, curriculum arrangement, teaching quality evaluation and other aspects, it can provide scientific and reasonable basis for higher vocational colleges to make decisions. For example, by analyzing data such as student performance and employment situation, teaching plans and curriculum Settings can be adjusted in time to improve the quality of education.

The scientific research that higher vocational colleges rely on also needs a large amount of experimental data support. The efficiency and quality of scientific research can be improved by managing the data of laboratories and scientific research projects. At the same time, data governance is also needed for the management and sharing of scientific research results in order to better promote academic exchanges and cooperation.

ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

In terms of research progress, at present, some scholars have conducted in-depth discussions on data governance in higher vocational colleges. Some researchers have put forward some data governance methods and frameworks, and applied them in practical scenarios. For example, through the construction of data warehouse, the establishment of data quality assessment model and other ways to achieve the goal of data governance in higher vocational colleges. In addition, some scholars pay attention to issues such as privacy protection and information security, and strengthen the protection of sensitive information in the process of data governance[1].

It can be seen that data governance in higher vocational colleges is of great importance from the perspective of big data, and certain research progress has been made in the academic circle. In the future, it is necessary to continue to deepen the research on data governance in higher vocational colleges, explore more effective methods and technologies, and provide better data support and decision-making basis for higher vocational colleges.

3. Build data governance model of higher vocational colleges from the perspective of big data

The establishment of data governance model in higher vocational colleges is an important work, which helps higher vocational colleges to effectively manage and utilize data resources. This research takes CIO (Chief Information Officer) system and master data management as the starting point, and initially builds the data governance model of higher vocational colleges from the perspective of big data. The simplification steps are as follows:

First, Clear goals: First of all, it is necessary to clarify the goals and needs of data governance in higher vocational colleges. This may include improving data quality, strengthening data security, optimizing data processes and so on.

Second, Formulate strategies: Based on clear goals, formulate data governance strategies suitable for higher vocational colleges. This includes determining the organizational structure, formulating policies and norms, establishing corresponding processes and process control, etc.

Third, Collection and collation: collect and collate all kinds of data resources within higher vocational colleges, and classify, clean and standardize them. Ensure the reliability and consistency of data.

Fourth, Establish a metadata management system: Establish a metadata management system for recording and managing all relevant metadata information. This information includes data definition, source, update frequency, department, etc.

Fifth, Develop access control mechanisms: In order to protect sensitive information and ensure compliance, access control mechanisms need to be established. Through role assignment and permission Settings, different users' access rights to different types of data resources are restricted.

Sixth, Implement data quality management: Establish a data quality management mechanism, including data verification, error correction and monitoring. Ensure the accuracy, integrity and consistency of data.

Seventh, Establish a risk management mechanism: Identify and assess potential data governance risks and establish a corresponding risk management mechanism. Discover and solve possible problems in time to ensure the smooth progress of data governance in higher vocational colleges.

Eighth, Continuous improvement and monitoring: Data governance is a continuous process that requires continuous improvement and monitoring. Periodically review and evaluate the effectiveness of the data governance model, and adjust and optimize it according to the actual situation.

Based on the rational implementation and planning of the above steps, this study gives the overall framework of the model from five dimensions: data source, big data pool, data warehouse, data mart and data visualization, as shown in Figure 1, aiming to help higher vocational colleges better manage and utilize their data resources, and improve the quality and efficiency of education and teaching.

ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

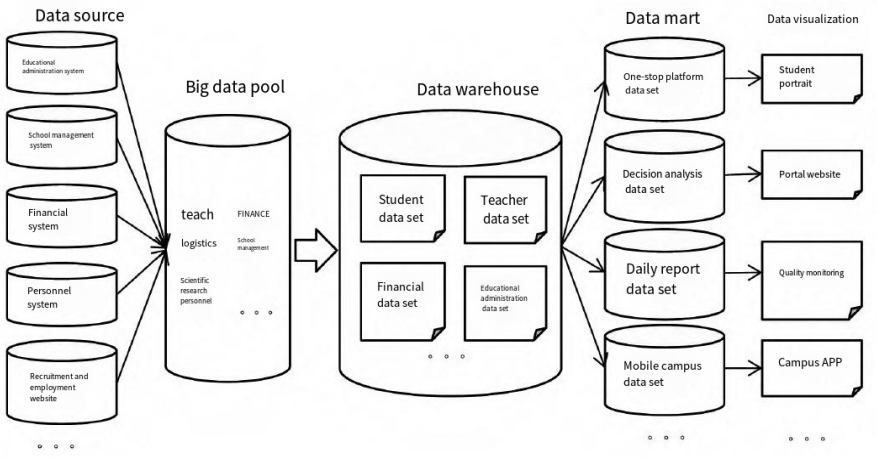

Figure 1: Data governance model of higher vocational colleges

On this basis, this study empowers the model and uses the ARCS motivation strategy model for system optimization, that is, improving the system governance system of higher vocational colleges for data acquisition, data storage, data processing and data exchange and other processes, so as to improve the management and utilization efficiency of data in higher vocational colleges, as shown in Figure 2.

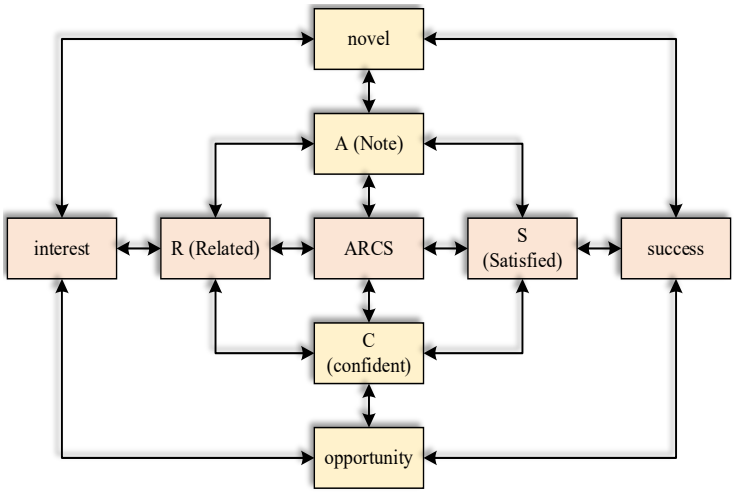

Figure 2: ARCS model optimizes the data governance system of higher vocational colleges

First, in terms of data acquisition, higher vocational colleges need to determine which data to collect and select appropriate methods and tools for collection. This may include extracting important data such as student information and course information from internal systems such as student management system and educational administration system, as well as obtaining external environment-related data through questionnaires, field observation and other ways.

Second, in terms of data storage, higher vocational colleges need to establish a suitable database or data warehouse to store all kinds of data collected. This can ensure the security and integrity of the data, and provide convenience for the subsequent analysis and application.

Third, in terms of data processing, vocational colleges need to clean, transform and integrate the original data collected. The cleaning process can remove erroneous or incomplete records, the conversion process can unify the data in different formats or codes, and the integration process can summarize the relevant data from multiple sources or departments.

Fourth, in terms of data exchange, higher vocational colleges need to pass on the processed and integrated data to relevant departments or decision makers to support their work and decision-making. This can be achieved by generating reports, making visual charts, and providing API interfaces.

After system optimization, feature extraction and evaluation verification of the model were carried out, and the correlation curve between the optimization times m and the accuracy of the model was drawn, and Table 1 and Figure 3 were obtained.

ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

Table 1: Relationship between optimization times m and model accuracy

m	A	B
1	50.01%	47.89%
2	79.45%	80.01%
3	89.45%	82.79%
4	95.57%	82.23%
5	98.34%	90.01%
6	98.89%	85.57%
7	100%	86.68%
8	100%	86.62%
9	100%	85.57%
10	100%	84.45%
Note: A in the table represents the training accuracy in the training data, and B represents the test accuracy in the test data		

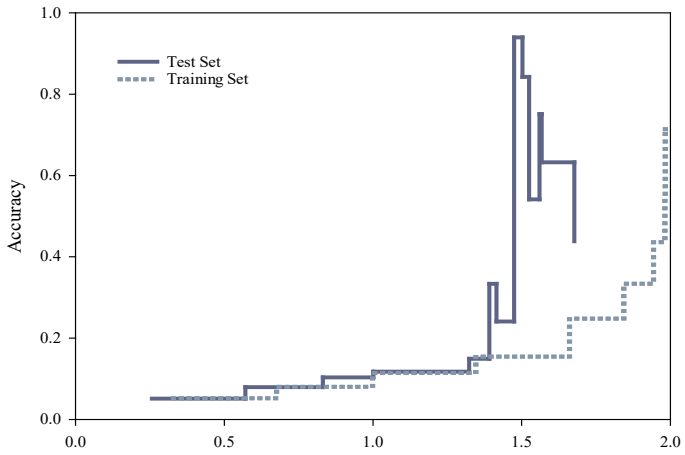

Figure 3: Optimization times m are highly correlated with model accuracy

As can be seen from the chart, when the optimization times m=7, the correct rate of the training set has reached 100%. It can be considered that the model is highly similar to the real data use, extraction, cleaning, optimization and other aspects, which means that the basic model proposed in this study is in line with the real situation, and practical simulation analysis can be carried out according to the model.

4. Problem-oriented, the actual situation of data governance in higher vocational colleges

With the rapid development of information technology, higher vocational colleges are faced with the challenges and opportunities of massive data. Data governance has become an urgent problem for administrators of higher vocational colleges. Based on the data governance model of higher vocational colleges constructed in this study, the following practical problems are proposed:

4.1 Heterogeneous data needs to be solved urgently

First of all, higher vocational colleges are faced with the management of a large number of heterogeneous data. These data, including student information, teacher information, curriculum information, etc., come from diverse sources and in different formats. Therefore, how to integrate and standardize these heterogeneous data has become an important issue.

4.2 Data privacy needs attention

Higher vocational colleges also face certain challenges in data security and privacy protection. Students' and teachers' personal information needs to be protected from leakage and misuse. At the same time, how to ensure data security when sharing and exchanging data is also an urgent issue.

ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

4.3 Data management needs to be improved

On the one hand, due to the large number of data sources and complex collection process, inaccuracies, incompleteness and even errors often occur in practical applications. Therefore, it is an important task to improve data quality and ensure data accuracy. On the other hand, higher vocational colleges also need to establish a sound data governance mechanism. This includes formulating relevant policies and norms, clarifying data management responsibilities and authority, and establishing data governance teams, among others. Only by establishing a scientific and reasonable data governance mechanism can we effectively improve the level of data management in higher vocational colleges[2].

5. Guided by governance, the optimization path of data governance in higher vocational colleges

5.1 Solving heterogeneous data problems

Through the actual parameter setting of the data governance model in higher vocational colleges, it is found that the objective problems of difficult management, access and use of heterogeneous data in higher vocational colleges can be greatly alleviated from the following aspects. First, data standardization, through the formulation of a unified data format and standards, so that the data between different higher vocational colleges can be compatible and exchange. This can facilitate the integration and sharing between different systems and improve the efficiency of data utilization. The second is to adopt the data integration scheme, the use of modern information technology means, the data scattered in each system to integrate and integrate. We can use ETL (Extract, transform, load) tool or develop custom interface to realize data integration, so as to eliminate the impact of heterogeneity. Third, the data should be cleaned regularly. Due to the problems of errors, redundancy or loss in different systems, the data needs to be cleaned and repaired. By using data cleaning tools or writing custom scripts for standardized processing, data quality can be improved and errors can be reduced. Using the model constructed in Chapter 2 of this study, we preliminarily simulated the governance of heterogeneous data after data standardization, as shown in Figure 4.

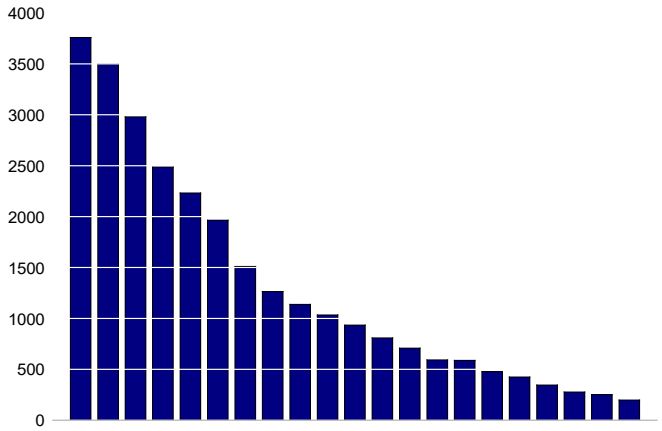

Figure 4: Influence of data standardization times on heterogeneous data quantity

As can be seen from the figure, after data standardization, the amount of heterogeneous data in the data warehouse of higher vocational colleges is significantly reduced. It can be seen that we should pay attention to the "disharmonious data" in the data governance system, and adopt standardization, cleaning or integration methods to eliminate the impact as much as possible, and convert data from different sources into the same format, such as unified date format and personnel number format. In order to facilitate subsequent analysis and application.

5.2 Solving data privacy issues

With the rapid development of data governance system in higher vocational colleges, the problem of data privacy becomes more and more prominent. In order to protect the personal information security of students and staff, it is imperative to solve the problem of data privacy in vocational colleges.

ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

On the one hand, higher vocational colleges should establish a sound data privacy protection system. The system should include clear policies, norms and processes to ensure that the collection, storage, processing and sharing of personal data comply with relevant laws and regulations. At the same time, data privacy awareness training for staff and students needs to be strengthened to raise their awareness of personal information protection. For example, higher vocational colleges can adopt technical measures to enhance data security. For example, advanced identity authentication technologies and access control mechanisms can be introduced to restrict access to certain sensitive information to only authorized personnel. In addition, encryption technology is used during data transmission to ensure that data is not stolen or tampered with during transmission.

On the other hand, higher vocational colleges can also cooperate with professional third-party institutions to strictly supervise the data processing process. These institutions should have professional competence and good reputation, and have rich experience in dealing with data privacy issues of similar institutions. By cooperating with these institutions, higher vocational colleges can reduce the risk of data leakage, improve the level of data privacy protection, and establish sound mechanisms for data security risk assessment and emergency plans. Regularly conduct data security risk assessment, identify potential security risks and take appropriate measures to solve them. At the same time, the development of perfect emergency plans, in order to be subjected to data leakage or other security incidents can be timely and effective response, only through the implementation of these comprehensive measures, in order to better protect the security of personal information in higher vocational colleges, and provide reliable guarantee for data governance and education development.

5.3 Improve the data management ability of higher vocational colleges

On the one hand, higher vocational colleges can improve their data management ability by establishing a sound data management system. This includes establishing unified standards for data collection, storage and processing, formulating clear data management policies and processes, and equipping professionals to take charge of data management work. At the same time, with the help of advanced information technology means, such as cloud computing and big data analysis, higher vocational colleges can better support the sorting, analysis and mining of massive data. On the other hand, higher vocational colleges should strengthen the construction of teachers and pay attention to the cultivation of knowledge and skills related to data management in teacher training. In this way, teachers' ability to collect, organize and use data can be improved, so that they can better use data to support teaching work and carry out scientific research projects[3].

In addition, higher vocational colleges should actively cooperate with enterprises to learn from their data management experience and technical means. Through cooperation with enterprises, higher vocational colleges can learn advanced data management concepts and methods, improve their own data management level, and apply it to teaching, scientific research and management. For example, in classroom teaching, teachers' teaching quality can be evaluated by students' course selection, exam scores and other indicators; In practice, students' comprehensive ability can be evaluated by indicators such as internship reports and project results. Through in-depth mining and analysis of these data, personalized guidance can be provided for teachers and personalized learning paths can be provided for students. In terms of scientific research, higher vocational colleges can use big data technology to mine resources such as academic literature and experimental data to improve the efficiency and quality of scientific research. Through the establishment of academic cooperation networks and knowledge maps, the sharing and exchange of scientific research results can be promoted and the development of disciplines can be promoted. In terms of student services, higher vocational colleges can make use of big data analysis technology to conduct personalized counseling and management for students. By analyzing data such as students' course selection, exam scores and social media behavior, targeted guidance and support can be provided to students to help them better adapt to campus life.

6. Conclusion

The results of data governance in higher vocational colleges are an important work under the vision of big data. With the rapid development and application of information technology, higher vocational colleges are faced with massive data processing and management demands. Through effective data governance, higher vocational colleges can make full use of and manage these data resources, and improve the efficiency of teaching quality, scientific research level and student services. In this study, the data governance model of higher vocational colleges optimized by ARCS model is innovatively

ISSN 2522-6398 Vol. 6, Issue 21: 112-118, DOI: 10.25236/FER.2023.062119

constructed from the perspective of big data. Through the accuracy analysis of the model, the optimization strategies for solving heterogeneous data problems, solving data privacy problems and improving data management ability are proposed in combination with the realistic and efficient data governance situation, aiming at ensuring the sharing and exchange of data within colleges and universities. In addition, decision support and intelligent analysis based on shared data are provided.

Acknowledgement

Project Name: Application and Research of Data Governance Results in Higher Vocational Colleges under the Background of Education Informatization, Project Number: SCHYB-2023-34.

References

[1] Songyuan Gao. Higher vocational colleges under the big data management ability to ascend to explore [J]. Journal of cooperation in economy and technology, 2023 (19): 124-125. The DOI: 10.13665 / j.carol carroll nki hzjjykj. 2023.19.045.
[2] Zhang Qin. Data Governance of Industrial College of Higher Vocational Colleges: Realistic motivation, basic ideas and promotion path [J]. Vocational Education Communication, 2023(01):29-35.
[3] Huang Youfu and Huang Zhongwei. Research on Data Governance in Higher Vocational Colleges under the background of "Double High Construction" [J]. Educational Information Technology, 2022 (04): 40-44.