Global Journal of Educational Studies
ISSN 2377-3936
2025, Vol. 11, No. 2

Data Management and Solutions at the Open University from a Big Data Analysis Perspective

Yufan Lin

Information Centre, Zhejiang Open University

Hangzhou, China

Tel: 086-177-9456-1560 E-mail: LYF19930811@126.com

Xintian Bu (Corresponding author)

Learning Resources Center, Zhejiang Open University

Hangzhou, China

Tel: 086-177-5717-0572 E-mail: buxintian@163.com

Received: September 17, 2025 Accepted: November 11, 2025 Published: November 20, 2025

doi:10.5296/gjes.v11i2.23363 URL: https://doi.org/10.5296/gjes.v11i2.23363

Abstract

From the perspective of big data, combined with the context of open education in universities, several key findings emerge from big data analysis in open education. First, it is essential to break down information barriers through relevant information technologies to achieve genuine campus openness. Second, universities should provide relevant data and information to support the development of an open university. This includes exploring the applications of big data analytics in open education and its extended uses in teaching services, campus services, and campus management. Third, data sharing and reuse should be facilitated through the construction of mutual platforms. Furthermore, hierarchical management should be implemented alongside intelligent applications, strengthened by top-level design. Pilot studies on Smart Alliance Campuses—which refer to campuses closely integrated with smart machines and software—should be initiated. Additionally, protecting data privacy and enhancing information security are critical. Overall, these findings highlight the need for comprehensive governance of big data in education.

The solutions include:

●Platform interoperability, data sharing, and reuse.

●Hierarchical management and intelligent applications.

●Enhancing top-level design through smart alliance campuses.

●Protecting data, maintaining privacy, and enhancing information security.

●Integrating big data governance into education.

Keywords: Data analysis, Open university, Big data

1. Background

Big data technology has emerged in recent years and is now deeply integrated with other enabling technologies such as cloud computing, artificial intelligence, and related fields. The extraction and analysis of massive volumes of data have become standard practices in the digital age. The prevailing consensus is that those skilled in data have the greatest potential for advancement. However, big data analysis techniques are still being tested and explored within colleges and universities. The convergence of open learning and research represents a novel development. Project development at higher education institutions increasingly depends on expanding big data platforms, and the transition from traditional data center platforms to big data platforms is an inevitable trend. To build the infrastructure necessary for colleges and universities to thrive in the long term, extensive data analysis must be employed as a technological foundation.

The theoretical basis can be analyzed as follows: The integration of big data technology into academic environments requires a robust theoretical framework to ensure its effective application. This involves understanding the fundamental principles of data management, including data collection, storage, processing, and visualization. Furthermore, the ability to interpret complex datasets is becoming an essential skill for both educators and students. The shift toward data-driven decision-making in higher education is supported by advancements in computational power and algorithmic development. Together, these elements contribute to creating an ecosystem where research and learning are enhanced through the strategic use of large-scale data resources.

It is essential to dismantle information silos using appropriate information technology to foster a genuine sense of campus openness. Promoting the development of an open campus also requires providing relevant data and information. By leveraging big data, we explore its application in creating an open educational environment within the university, focusing on teaching services, campus services, campus management, and other areas. This investigation examines how extensive data analysis supports open education and enhances the diverse forms of knowledge demonstrated in the learning and work of teachers, students, and other campus users.

2. Research Methodology

Following are the Research stages and Research methods and Research tools used at these stages:

2.1 Research Stages

● Preparation Phase: Define the project's objectives and develop a comprehensive plan. This plan should include the project's scope, budget, timeline, costs, and resource allocation.

● Implementation Phase: The project team executes the tasks according to the plan and develops an open data platform for the Open University.

● Summarization Phase: Summarize and evaluate the project's implementation process, and prepare relevant reports and papers.

2.2 Research Tools

● Data collection tools: e.g., questionnaires (e.g., Zotero, Paperpile, Dovetail, EndNote), interview guides, observation forms, and experimental equipment.

● Data analysis tools: e.g., statistical software (SPSS, R, SAS), content analysis software, qualitative analysis software (NVivo, Atlas.ti).

● Literature management tools, such as Zotero and EndNote, assist researchers in organizing and citing literature.

● Research tools: e.g., SurveyMonkey and Typeform, used for designing and distributing online surveys.

3. Data Sources and Sample Characteristics

Following are the data sources and sample characteristics at the Open University from the big data perspective. Our organization is currently developing a lifelong education big data warehouse that encompasses data ingestion, detailed processing, aggregation, and application layers. The data middle platform integrates 729 tables from 21 business systems, with a total data volume of 1.25 TB. We have defined 10 lifelong education data domains and 8 core thematic domains. At present, we have preliminarily constructed the big data warehouse, refined data standards, established an asset catalog, and implemented management systems to ensure secure data sharing.

Most of the data comes from the Lifelong Education Big Data Warehouse. The sample includes a diverse range of learners from various educational backgrounds, age groups, and geographic locations. It represents a mix of traditional students, working professionals, and adult learners pursuing continuing education. This diversity ensures that the data provide a comprehensive view of lifelong learning patterns and behaviors. Additionally, the temporal aspect of the data captures longitudinal trends, allowing for the analysis of how learner engagement and outcomes evolve. These characteristics make the dataset particularly valuable for identifying key factors influencing educational success across different demographics. The sample characteristics reveal a balanced distribution across genders, with a slight predominance of female learners. The age distribution indicates a significant proportion of participants aged between 25 and 45, reflecting the active engagement of working professionals. Educational backgrounds vary widely, ranging from high school graduates to individuals holding advanced degrees, which enhances the dataset's heterogeneity. Geographically, the sample encompasses both urban and rural areas, capturing regional differences in learning preferences and resource availability. Furthermore, the inclusion of various learning modalities—such as online courses, blended learning, and traditional classroom settings—adds layer of depth. This multifaceted sample enables robust analysis, allowing researchers to explore correlations between demographic factors and educational outcomes while accounting for contextual variables that may influence the results.

4. Problems

4.1 The Problem of Data Silos

According to Wu (2020), a data silo refers to data within each department functioning as an isolated island, making it difficult or impossible to connect and interact with data from other internal departments. The development of university informatization aligns with the mainstream construction framework of "one network, three channels, and two platforms," as outlined in the National Education Informatization "13th Five-Year Plan" and the "Education Information Technology Development Plan." Specifically, initiatives such as "Smart Alliance Campuses" (including network integration), management information bases (featuring extensive data analysis and decision support systems), learning resource bases, and other related products and services have become central to this unified construction approach. This approach reflects the homogeneity of core operations and highlights the need for further research and development of peripheral products. Due to the absence of top-level design in planning and the siloed construction issues stemming from "Education Informatization 1.0," data silos have become increasingly prevalent. Promoting deep data mining to support decision-making—an outstanding aspect of university information management beyond routine data maintenance—requires significant effort.

4.2 Personalized, Widespread, and Multi-terminal Learner Group Characteristics

The Open University supports customized, distributed, and multi-terminal student groups due to its emphasis on open education. The Open Institution—a new type of university built on the Internet and utilizing multiple media technologies as platforms, primarily for adult learners—has entered the 2.0 phase with the expansion of information media (Bao, 2018). For example, Zhejiang Open University has adapted to the integration of online and offline learning styles by leveraging resources in its learning resource library. The Zhejiang Learning Network, based on a B/S architecture, is provided to students; the Zhejiang Learning app, based on a C/S architecture, offers convenient mobile access, enabling learners to study anytime and anywhere; and Microtelecom University, a WeChat-based platform, allows learners to log in via WeChat for learning and information inquiries. It is essential to carefully adapt to popular learning styles such as socialized and personalized learning, as well as to coordinate and manage learning channels across each platform and system, so that students can fully benefit from the convenience of multifaceted learning options.

4.3 The Big Data Cult and the Danger of "Big Data Only"

According to research, Big Data is characterized by the 4Vs: volume, velocity, variety, and value (Nan, 2020). The risk of relying solely on "big data for individual users or the entire population may arise from blind worship and excessive use of big data. Big data is delivered to each individual through several processes, including data cleansing based on massive data collection, data mining, analysis of valid data, extraction of relevant information, and the provision of adaptive services. The direction of extensive data mining and analysis is heavily influenced by learning and inferring the characteristics of individual users. However, this approach is vulnerable to issues such as uncritically pushing similar and identical user

interest points, which can lead to diminishing marginal utility.

4.4 Data Security Concerns

The fluid and interchangeable nature of big data maximizes information sharing; however, indiscriminate data mining, excessive data collection, and misuse lead to numerous data security issues (Jin, 2020). In higher education, the campus Wi-Fi network and its associated infrastructure—such as wireless hotspots and network access points—are the most prominent examples. Access to university intranet office websites and on-campus informational repositories, including knowledge networks, often lacks additional firewalls and other security measures, rendering them highly vulnerable to hacker attacks. For instance, the alteration of images on Zhejiang University's official homepage has been exploited to disseminate illegal propaganda. Furthermore, when university data is shared with external organizations, such as suppliers, it is frequently transmitted without encryption or via insecure channels, increasing the risk of leakage of sensitive student and parent information and posing a significant threat to internet privacy security.

5. Solutions

The solutions can be proposed as follows:

5.1 Platform Interoperability, Data Sharing, and Reuse

Construct an open data platform using open-source software to facilitate data sharing and reuse while ensuring platform compatibility. We are expanding this initiative to develop university platforms, specifically creating a platform for public university data and collecting data through API (Application Programming Interface) technology. Data integration, sharing, and the application of business systems are further enhanced by combining data collectors to achieve efficient data storage, interchange, and sharing. We aggregate, filter, and categorize a substantial amount of video teaching materials related to widely used MOOC and SPOC course systems, as well as technological tools such as micro-classes and flipped classrooms, to build the province's university application data resource base and effectively create a blended learning environment for students. The Zhejiang Provincial Education Bureau, relying on the Zhejiang Provincial Education Technology Center, is constructing a unified provincial database from the ground up. This database provides universities in Zhejiang Province with a standardized reference audit table structure, specifies mandatory fields and reference field classifications for primary tables, and offers essential construction information, including bilingual (Chinese and English) cross-reference annexes. From a more visual perspective, as illustrated in Fig. 1 below, the Zhejiang Learning Code can count the number of applicants in real time and map this data onto a visual dashboard, thereby producing a graphical representation of the analysis results and achieving the goal of data sharing. The Zhejiang Learning Network is accessible to both academic and non-academic educational institutions.

Figure 1. Zhexuema running data visualization screen

5.2 Hierarchical Management and Intelligent Applications

The collected sample data can be intelligently utilized through artificial intelligence (AI), augmented reality (AR), virtual reality (VR), and other advanced technologies to create immersive experiences, support personalized learning, and cater to the needs of diverse, ubiquitous, and multi-platform learner groups. A notable example demonstrating the integration of AI, extensive data analysis, new media technology, and data visualization is the Zhejiang Kai Le Learning Port located on Jiaogong Road in Hangzhou, Zhejiang Province. Starting with hierarchical administration, matching algorithms can be employed to manage and monitor adaptive learning, while interaction modules can be used to develop an advanced adaptive online education system.

5.3 Enhancing Top-Level Design through Smart Alliance Campuses

In August 2017, the Ministry of Industry and Information Technology (MIIT) issued a three-year action plan (2018–2020) to support the development of a new generation of the AI industry. Building on this, the Ministry of Education formulated the “Development Plan for AI+Education” for the Open University, using relevant data and the university’s specific circumstances as a foundation. This plan emphasizes the importance of focusing on the top-level design of “AI+Education” and conducting pilot studies for the Smart Alliance Campus. The Smart Alliance Campus concept and architecture aim to enhance the precision of educational services. The implementation process should fully leverage big data technologies to enable dynamic and extensive data collection, real-time analysis, and decision support. It should also integrate emerging technologies such as the Internet of Things, cloud computing, mobile 5G, and connect to various teaching applications. Additionally, establishing a robust network security assurance system is critical. This system must facilitate real-time monitoring of vulnerability detection and security testing, as well as effectively prevent network security threats during holidays. Figure 2 illustrates the topology of the

university network under construction, including the deployment of top-level security configurations, monitoring through a security situational awareness system, attention to network attacks, log auditing and inspection systems, review of logs from each security device, and focus on database firewalls, fortress machines, and antivirus software. It also involves analyzing the network security architecture, troubleshooting security risks, and implementing necessary measures.

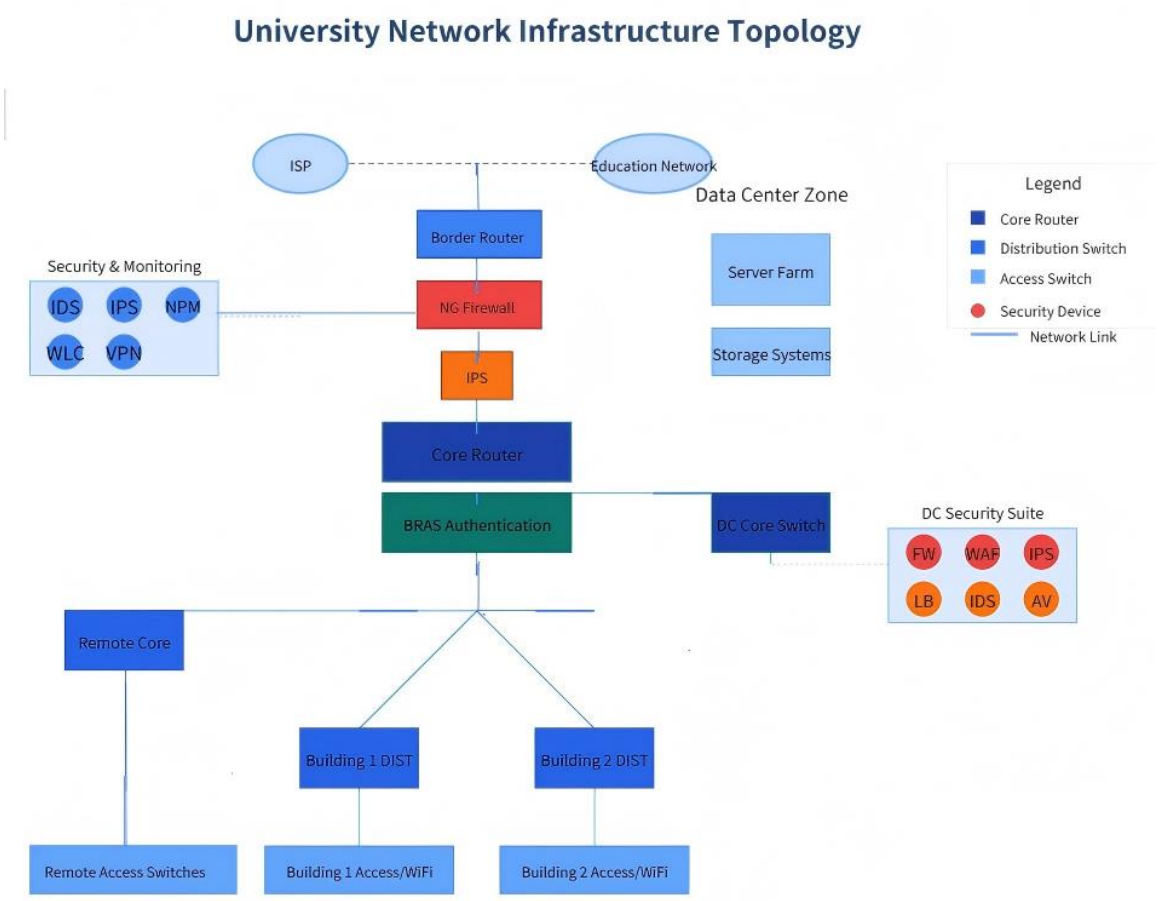

Figure 2. University network infrastructure topology

5.4 Protecting Data, Maintaining Privacy, and Enhancing Information Security

The school should implement information security techniques and sustainable management practices to protect data and privacy. These measures are essential to address various data security challenges arising from data mining, excessive data collection, and misuse. Key laws and regulations, such as the Network Security Law and the Implementing Rules of Equal Protection 2.0, have been enacted, providing crucial policy support for effective information security efforts at the governmental level. The rise in network attacks has indirectly accelerated the development of information security solutions, enabling colleges to select from a wider range of products with enhanced features. Given the significant data integration involved at the Open University, it is recommended to establish an electronic file data management platform separate from the school's main data center. This platform would allow

electronic file data to be maintained, stored, archived, and utilized efficiently and independently. During the routine maintenance and processing of big data, the school will implement firewalls and other security barriers, while security tools such as fortress machines will log information accessed by system administrators and users to facilitate supervision.

The Open University will exercise caution when sharing personal information with parties other than suppliers to protect student data privacy and prevent unauthorized disclosures. It is recommended that universities employ technologies such as tokens and passwords, along with processes like restricting access rights, to better safeguard user privacy and enhance information security.

5.5 Integrating Big Data Governance into Education

The four major components of integrated big data governance in education are outlined below. First, we examine the data collected by sensing devices and intelligent sensor technologies, including RFID and GPS, to understand data-gathering patterns at the physical level. The extensive educational data resources collected are then carefully analyzed using three essential data operations: data mining, data analysis, and storage. Data mining techniques cover a variety of areas, including decision simulation, sentiment analysis, and learning prediction. Modern data analysis methods, grounded in comprehensive data examination, are employed for statistical, graphical, and predictive analyses. Visualization analysis, valued for its effectiveness, is utilized by the Open University. Visualization displays can be implemented on campus to provide real-time updates on students, course materials, and learning outcomes.

This text discusses the development of the Leader Driving Bin, a tool designed to visually display the results of extensive campus data analyses. These analyses cover various aspects, including teacher attendance, stranger behavior, classroom and meeting room utilization, and energy consumption for air conditioning and lighting. The goal is to explore the creation of a new learning environment characterized by connectivity, transparency, and the exchange of excellent instructors. We advocate for the establishment of an open campus that leverages data and information sharing. The discussion will focus on the application of extensive data analysis in open education within a university setting, examining how it can enhance the work and learning experiences of teachers, students, and audience members. Additionally, we will consider its use in teaching services, campus operations, and other related areas.

6. Conclusion

We aim to create a new learning environment that emphasizes connectedness, openness, and the exchange of expertise among top educators. Our commitment is to advance big data analytics to make it intelligent, pervasive, and better aligned with the increasingly transparent and collaborative nature of modern organizations. We provide information and data support to foster an open campus environment. The application of extensive data analysis in open education will be demonstrated within a university setting. Additionally, we will explore its potential uses in teaching services, campus operations, administration, and other areas, as well as how it can support the work and learning of instructors, students, and other

stakeholders. However, limitations and suggestions for future research should not be overlooked. For instance, the risks associated with relying exclusively on data are rapidly increasing. Analysts may lose direction within the Lifelong Education Big Data Warehouse and become overly influenced by the information provided by big data. Another significant concern is data security, particularly during the data storage process. The most sensitive data parameters—including ID numbers, bank account numbers, and private telephone numbers—are at a heightened risk of being stolen for illegal purposes, potentially leading to serious consequences. Suggestions for further work include enhancing the transparency of data usage and establishing stricter regulations to protect sensitive information. It is also crucial to develop more advanced encryption technologies and multi-layered security protocols to minimize vulnerabilities. Additionally, fostering interdisciplinary collaboration could provide new perspectives and innovative solutions for addressing these challenges. Future research should focus on balancing the benefits of big data with the need for privacy and ethical considerations, ensuring that technological advancements do not come at the expense of individual rights. By doing so, the field of lifelong education can continue to evolve sustainably and responsibly.

Acknowledgments

We greatly appreciate the valuable contributions of our community advisory committee members. We would also like to thank the Zhejiang Open University and every team member who took the time to participate in this study.

This work was supported by Zhejiang Open University General Project “Research on the Construction and Operation Mechanism of Inclusive Education Public Service System from the Perspective of Human Capital” [project number XKT2023Y17]; Zhejiang Open University General Project “Innovative Models and Implementation Pathways for AI-Enabled Smart Party Building and Integrated Online Ideological and Political Education in Higher Education Institutions”[project number 2025xkt03]; 312 Talent Training Project of Zhejiang Open University; Zhejiang Provincial Education Science Planning Party Building Special (Colleges and Universities) Project “Innovation and Practice of Intelligent Party Building and Network Civic and Political Integration Mode in Colleges & Universities Empowered by Digitization” [project number 2024DJG017]; Zhejiang Provincial Archives Science &Technology Project “Exploring the Construction Path of College Archives Wisdom Service Platform Based on the Threshold of Data Co-construction, Sharing & Integration Development”[project number 2024-51].

References

Bao, R. (2018). The Change of Teaching and Learning in the Era of Artificial Intelligence and the New Exploration of Open University 2.0. *Journal of Distance Education,* 36(03), 25-33. (In Chinese). https://doi.org/10.3969/j.issn.1672-0008.2018.03.003

Jin, Y. (2020). Research on Personal Information Security from the Perspective of Big Data. *Legal Expo*, (1), 15-17. (In Chinese) https://doi.org/10.3969/j.issn.1005-1595.2020.01.005

Nan, G., & Xia, Y. (2020). Educational Science Research in the Era of Big Data: Possibilities,

Risks and Strategies. *Journal of Educational Science of Hunan Normal University, 19*(04), 87-94. (In Chinese) https://doi.org/10.3969/j.issn.1674-831X.2020.04.010

Wang, Y. G., Xu, J. Q., & Ding, J. H. (2020). Education 4.0 Global Framework: Future Schooling and Paradigm Shift—Interpretation of the Report of the World Economic Forum "Schools of the Future: Defining a New Education Model for the Fourth Industrial Revolution". *Journal of Distance Education, 38*(03), 3-14. (In Chinese) https://doi.org/10.3969/j.issn.1672-0008.2020.03.001

Wu, Y. (2020). City "Wisdom" Must Open Up Data Silos. *Nanjing Daily*, 12-24. (In Chinese) https://doi.org/10.3969/j.issn.1006-9670.2020.05.002

Copyright Disclaimer