Design of Digitalization Application for Hospital's Heterogeneous Business Data in the Ai Era

Gao Yu[1,a], Xunwei Lu[2,b]

[1] MBA Center, School of Management, University of Science and Technology of China,Hefei, Anhui, China,230000;

[2] School of Public Health, Fudan University, Shanghai, China,20032.

[a] yugao1986@ustc.edu, [b] ellenlu@189.cn

Abstract. The purpose of this article is to construct a hospital big data platform to achieve centralized management of the heterogeneous business systems and historical business data in use by the hospital. By integrating multiple business systems in the hospital through the hospital big data platform, data standards are established and a unified hospital data model is constructed to compatibility with heterogeneous diagnostic data information. This enables standardized storage and data analysis of both new and old business data, providing accurate, scientific, and multi-dimensional data support for clinical applications, scientific research and teaching, operation and management, and intelligent assistance decision-making in hospitals.

Keywords: medical big data, big data platform, heterogeneous business.

1. Introduction

In recent years, the Communist Party of China (CPC) Central Committee and the State Council have issued relevant documents such as the "Healthy China 2030 Planning Outline," the "Action Plan for Promoting Big Data Development" by the State Council, and the "Guiding Opinions of the State Council General Office on Promoting and Regulating the Development of Health and Medical Big Data Applications." These policies and guidelines have put forward requirements for the construction of smart hospitals at the macro design level from the top. Furthermore, the "Action Plan for Promoting High-quality Development of Public Hospitals (2021-2025)" further clarifies specific actions for the high-quality development of public hospitals during the "14th Five-Year Plan" period, including the construction of a high-level network of public hospitals and "three-in-one" smart hospitals. The long-term mechanism of hospital management in our country, which emphasizes "Promoting construction through evaluation, promoting management through evaluation, combining evaluation and reform, and focusing on content" plays an important role in the construction of smart hospitals [1]. The hierarchical evaluation of the functional application level of electronic medical record systems and the standardized maturity evaluation of hospital information interoperability impose different requirements on hospital data centers and big data platforms [2, 3]. Therefore, the construction of big data platforms is gradually becoming a standard configuration for smart hospital construction.

2. Research Background

Currently, hospital big data platforms are mainly used for research purposes. These platforms enable the collection and cleansing of business data, data governance and standardization, semi-structured processing, and big data search capabilities. By optimizing the research process, these platforms can improve the efficiency of scientific research [4-8]. The quality of research data is of utmost importance, and the success of a hospital big data platform depends heavily on the quality of the hospital's data. In China, medical data in hospitals is often dispersed across various business systems, with inconsistent data structures and standards [9]. The construction of a medical big data platform faces challenges such as numerous health information standards, difficulties in mining medical data, and security risks in data sharing and application [10].

In addition, the 2020 National Monitoring Report on the Performance Evaluation of public hospitals at the third level nationwide showed that the average level of application for electronic medical record systems in third-level public hospitals increased from 3.23 in 2019 to 3.65. In the context of the "three-in-one" smart hospital construction, leading healthcare IT vendors have proposed new hospital information systems based on the new era architecture,, such as iMedical by Donghua, WiNEX by Winning Health, and Esmart by Hangzhou Century Co., Ltd. The introduction of these new architecture-based hospital information systems has led to the replacement of existing HIS (Hospital Information System) and CIS (Clinical Information System) in many healthcare institutions. The upgrading and replacement of hospital information management systems pose new challenges to the construction of big data platforms in multiple hospitals: how to incorporate data from historical business systems into the platform for unified management and utilization.

3. Design Methodology

A big data platform needs to provide comprehensive data collection, construction, management, and utilization capabilities throughout the data lifecycle. It should significantly improve data governance levels and meet the diverse data-driven application needs of healthcare institutions, thereby strengthening their digital capabilities through data empowerment. To achieve this goal, the architecture of the big data platform consists of seven components: data collection, data transformation, data center, data governance, data services, intelligent operations, and data applications.

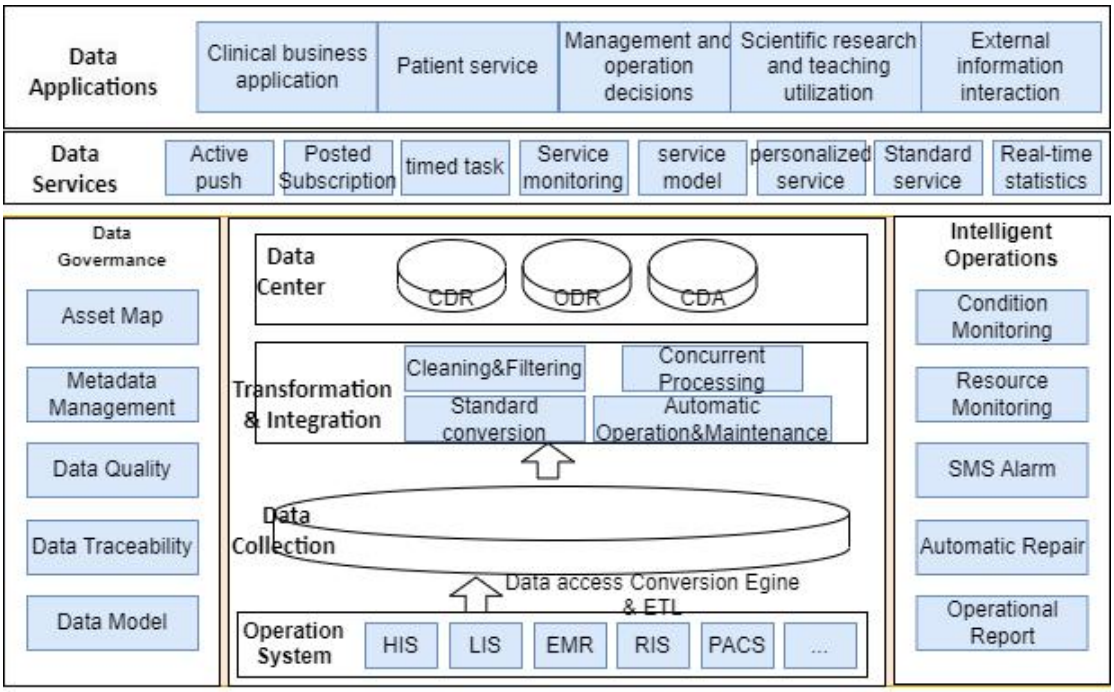

Figure 1 A Schematic diagram of the big data platform architecture for heterogeneous data-oriented hospitals

3.1 Data Collection

Data collection is the entrance for data connection to a big data platform. All data comes from various sources such as healthcare institution's business systems, logs, files, etc., which are scattered across different environments and storage platforms, making it challenging to utilize. The main function of data collection is to conveniently collect various heterogeneous vendors' and data sources' data into the big data platform for centralized storage, facilitating subsequent processing and modeling. For the replacement of existing business systems in hospitals, data collection is

mainly conducted using the Extract, Transform, Load (ETL) method. Depending on the timeliness of data access, there are offline batch aggregation and real-time collection methods. For systems in use or newly launched, database synchronization and message queues are commonly used for data collection.

3.2 Data Transformation and Integration

The data collected and aggregated to the platform through the data acquisition module is usually in its raw state and is not suitable for direct use in business operations. The data transformation engine provides a complete set of tools for data processing and process control. Experienced data developers and algorithm modelers can utilize the functionalities provided by the data transformation engine to quickly transform the data into a valuable format for business users.

For structured data, the objects of data transformation processing not only include associate operations between data from different business systems, but also include integrating data from different versions of business systems into a standard model. For unstructured data, such as electronic medical records, the data transformation module with its built-in NLP engine is required to structure the data to meet the high-level requirements for electronic medical records and the scientific research needs of hospitals.

3.3 Data Center

With the data collection and data transformation modules in place, the big data platform has acquired the basic capabilities of a traditional healthcare data center platform. It is now able to aggregate and develop various types of data, enabling the establishment of various themed data centers, such as clinical data centers and operational management data centers.

In order to ensure compatibility with the business needs of mainstream healthcare information vendors in China, the data center refers to relevant standards and specifications such as the Health Information Dataset, Data Elements, and Electronic Medical Record Dataset. It constructs more than 150 business models covering a wide range of topics, including patient information, patient services, medical orders and prescriptions, laboratory services, diagnostic services, blood transfusion services, nursing, medical records and documents, surgical anesthesia, and physical examinations, and others. The coverage is wide and deep, almost covering all the business needed by hospitals.

3.4 Data Governance

The data assets established through the data center are more technical in nature, making it difficult for business users to understand. Data governance aims to provide a better understanding of the data to medical professionals through ensuring data quality, showcasing data flows, and managing data assets in a way that is more comprehensible to them (while considering permissions and security control). Data governance management includes managing and displaying data asset catalogs, metadata, data quality, data lineage, data models, and other aspects, providing a more intuitive way to display hospital data assets and enhance enterprise data awareness.

3.5 Data Services

While data acquisition and data transformation modules help build data assets, and data governance showcases the data assets of healthcare institutions, the true value of healthcare business data remains untapped. Data services transform data into a service capability, enabling data to actively participate in healthcare operations and activate the entire data platform. The ability to rapidly generate services and establish a data services system encompassing service control, authentication, and metering is where the value of a data platform lies.

3.6 Intelligent Operations

Through the aforementioned modules such as data acquisition, data transformation, data center, data governance, and data services, the construction and development of the entire data middleware have been completed, and they have already delivered certain value in business operations. To ensure the platform's smooth operation, the platform provides operation and maintenance as well as security management modules. Operations and maintenance, along with security management, form the foundation for the healthy and continuous operation of the data middleware platform.

3.7 Data Applications

Based on various themed data centers, the platform provides data applications for clinical, operations management, data analysis, and research purposes. These applications include operational analysis, data reporting, and more. Leveraging a microservice architecture, data applications can be easily integrated and deployed on the data platform, allowing for flexible and customizable combinations.

4. Achievements in Construction

4.1 Integration of Heterogeneous Business Data

Isolated data islands cannot effectively support hospital operations decision-making or adapt to rapidly changing front-end business requirements. Medical institutions need to establish a data center that integrates data from these isolated sources, combining old and new data models to quickly form data service capabilities that support fine-grained management of the hospital.

By constructing a standardized data center, previously scattered in various campuses and business systems, data information can be standardized and integrated, thus improving the efficiency of the data center. This, in turn, provides accurate, scientific, and multidimensional data support for data analysis, clinical applications, research and teaching, operational management, and intelligent decision-making within the hospital. Particularly for medical institutions planning to replace their business systems, a standardized data center can effectively address concerns related to the utilization of historical data.

4.2 Expansion of Cross-System and Cross-Department Data Applications

In terms of data resource utilization, the construction of a big data platform enables the discovery of missing processes or weaknesses in clinical and operational management, supporting the identification and improvement of issues in hospital informationization projects. The establishment of a clinical data center facilitates the sharing and integration of clinical diagnosis and treatment data, enhancing the convenience of clinical work. By connecting to the national health data platform, data circulation is achieved from hospital to national regulatory agencies.

In terms of operations management, the big data platform aggregates various business data resources within the hospital. Operations management data can serve functional departments such as the operations management department, quality management department, and performance management office in their daily management work based on data analysis. Evaluating performance indicators in public hospitals based on the big data platform further clarifies the responsibilities and goals of various functional departments, strengthening the support of business data for their daily work. Data applications such as Mobile BI enable hospital leaders to access real-time operational information using mobile devices, supporting fine-grained management of the hospital.

4.3 Establishment of Enterprise-level Data Center Standards

In accordance with national standards and actual requirements, the data center for hospitals has formulated data standard specifications for the data platform. This includes data dictionaries and standards for medical and health service institutions, administrative management organizations,

healthcare professionals, and other basic data resources. The data center is built around the core of electronic medical record (EMR) data, with a database and platform services that support data mining, analysis, and sharing through data exchange capabilities within healthcare institutions. The data standards of the data center comply with national and healthcare industry information and data standards or specifications, following the latest research achievements in domestic and international health information standards, and meeting the relevant regulatory requirements of the country.

5. Discussion

The hospital big data platform has achieved an integrated, standardized, and patient-oriented collection of clinical data that is independent of specific medical procedures. It serves as a new layer of data storage structure centered around patients, supporting clinical diagnosis and treatment, teaching, and research activities. The construction of the big data platform integrates business data from different clinical information systems, stores clinical data across multiple domains, facilitates data transmission and sharing among various information systems in hospitals, and provides a comprehensive and refined data platform for management analysis. It effectively solves the problem of "information island " in healthcare institutions.

The big data platform serves as a data storage center while providing various data access services that comply with standards, legal norms, and security requirements for the hospital's business systems. It offers a unified, doctor-oriented integration and central display of clinical data for different healthcare applications, assisting in improving the quality of healthcare services. Based on the various services provided by the data center, scientific and intelligent data support can also be provided for clinical medical activities.

The "14th Five-Year Plan for Health Standardization Work" emphasizes the need to improve the standard system for health information, including the development of standards for basic, data, application, technology, management, and security and privacy categories. Building a standardized data center aligns with the goal of establishing a standard system that promotes the high-quality development of public hospitals, fosters data sharing and interoperability among medical institutions in different regions and at different levels, and contributes to enhancing the scientific, standardized, and refined level of hospital management in China.

References

[1] Cao, R. (2010). Continuous improvement of medical quality and patient safety. Chinese Hospital, 14(1), 2-6.

[2] National Health Commission Statistical Information Center. (2020). Evaluation scheme for the standardization maturity of hospital information interconnection. Retrieved from http://www.nhc.gov.cn/mohwsbwstjxxzx/s8553/202008/e80dafa1334c44c38f644602406a4973.shtml

[3] National Health Commission. (2018). Grading evaluation criteria for electronic medical record systems (trial implementation). Retrieved from http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=3cae6834a65d48e9bfd783f3c7d54745

[4] Yu, G. (2016). Design of clinical research application based on medical big data. Chinese Digital Medicine, 11(9), 15-17.

[5] Zhu, M. (2020). Research on clinical research big data platform. Chinese Digital Medicine, 15(7), 17-35.

[6] Zhang, L., & Yang, Z. (2020). Construction and application of medical research big data platform based on Hadoop. Chinese Digital Medicine, 15(10), 82-84.

[7] Tang, B., Yao, L., et al. (2021). Application practice exploration of hospital research big data platform. Chinese Digital Medicine, 16(11), 104-108.

[8] Zhang, Q., Huang, J., et al. (2022). Practice exploration of emergency special disease database construction based on unified hospital big data platform. Chinese Digital Medicine, 17(4), 19-25.

[9] Wu, Y., Li, M., et al. (2021). Construction and data governance practices of a hospital-level medical big data platform for clinical research. Journal of Medical Science and Research Management, 34(2), 81-86.

[10] Guo, Q., & Wang, C. (2021). Opportunities, challenges, and development of medical big data platform. Journal of Medical Informatics, 42(1), 2-8.