Design and Implementation of Data Middle Platform

Zheng Haotian
College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China, 349600418@qq.com, Corresponding author

Li Tao
College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China
litao@wust.edu.cn

Yan Song
College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China

ABSTRACT

Under the background of the continuous advancement of China's digital process, the digital transformation is entering a new stage, and the role of the Data Middle Platform is becoming more and more important. Data middle platform is to abstract the common needs of users and build them into platform-based and componentized system capabilities, which are shared with various business units in the form of interfaces, services, components, etc., so that enterprises can quickly and flexibly call resources to build solutions for specific problems and enable business innovation and iteration. The Data Middle Platform helps enterprises to achieve business interconnection, resource coordination and information sharing with unified standards and process specifications. The Data Middle Platform collects data from the background and the business center for data sharing and integration, management and service application, and provides quick decision response, fine operation and application support for the front desk business. In the traditional management system, there are many organizational departments with diversified business and data. The business and data between departments are independent and cannot communicate with each other, forming a "data island". A large number of repeated function construction leads to waste of resources. In view of this situation, this paper proposes the design and implementation of data platform based on management system to promote the standardization and intellectualization of management work.

CCS CONCEPTS

• **Information systems** → Information systems applications.

KEYWORDS

Data Middle Platform, management, share, resource, management

ACM Reference Format:
Zheng Haotian, Li Tao, and Yan Song. 2021. Design and Implementation of Data Middle Platform. In *2021 2nd International Conference on Artificial Intelligence and Information Systems (ICAIIS '21), May 28–30, 2021, Chongqing, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3469213.3470401

ICAIIS '21, May 28–30, 2021, Chongqing, China

ACM ISBN 978-1-4503-9020-0/21/05…$15.00
https://doi.org/10.1145/3469213.3470401

1 INTRODUCTION

As Internet popularity at home, big data, artificial intelligence, such as cloud computing technology rapid development, companies are also accompanied by the more diversified business, will generate huge amounts of data within the enterprise, and in the case, the traditional enterprise internal most interfaces, services, and database are developed independently, this situation led to excessive scattered data between various business systems, lead to the generation of the data island, has experienced a ferment, precipitation and long data, China has become the new tuyere, make enterprise internal between various business system data flow operation. Data Middle Platform, we are curious about its origin, but also hope to see the whole picture, through it to bring us a better future, "data to use", this phrase is the source of the rapid development of data in the platform is also the common goal we strive for. This paper first introduces the research background and significance of data in the platform, describes the relevant research on data in the platform at home and abroad, analyzes the characteristics and problems in the development process of data in the platform, and finally expounds the application of the concept of data in the management system.

2 RESEARCH BACKGROUND

Enter the era of big data. As a product of innovation and practice in China, data platform is inseparable from its originator, Alibaba. First of all, Alibaba's practical experience in data platform based on internal massive data applications, as well as its innovative attempt in data enabling business with Internet technology and thinking as the core, such as new retail and new finance, awaken the industry to comprehensively follow up and try the concept of "data platform". The goal of the platform is to provide universal data services. With the promotion of "Internet plus" action computing and "intelligence +", digital industrialization and industrial digitalization become the two foundations of digital economy. Different enterprises in different industries have different data application requirements at different stages. The construction of Data Middle Platform is a continuous improvement process. In this process, the data of different stages in different scenarios are constantly updated and iterated.

It is also an important trend to transform and optimize the data middle station through artificial intelligence technology. Through artificial intelligence technology, the data middle station can be more agile and efficient, and integrate artificial intelligence technology into the underlying tools and products of the data middle

station. The middle station can independently analyze the data and business dependence between business systems, and build intelligent recommendation data standards and data governance Discussion.

3 DESIGN OF DATA MIDDLE PLATFORM

The goal of building a Data Middle Platform is to keep the data in use. Through the methods, tools and operational mechanisms provided by the Data Middle Platform, data can be transformed into a service capability to make data more convenient for business use.

In the traditional management system, the problem of data island is quite serious. The accumulated data of the business system far exceeds the previous level of data, the business form and data become diversified, the organizational level is complex, the new business needs continue to emerge, and the concept of big data governance, value and so on have a new understanding. The Data Middle Platform platform is the result of the natural evolution of the digital transformation process of the management system, which is divided into the following scenarios

(1) The department keeps ledger, and the superior department supervises the execution of tasks by the subordinate department, and reports the completion of tasks to the superior department after statistics. Through the data platform, the system automatically distributes tasks to each department according to the elements of ledger. On this basis, the superior department can independently assign tasks to the subordinate management department according to individual needs and supervise the completion of tasks. The organs and units supervise and guide the three meetings and one lesson of each department and the development of organizational activities of each department, and report to the superiors one by one the annual organizational activities, three meetings and one lesson of each department, departmental ledger and personnel learning.

(2) Annual assessment: according to the rules set in the system, departments at all levels will be assessed according to the degree of completion of tasks, points will be scored according to the frequency of completion, and departments will be urged to speed up the progress of work according to the points.

(3) Personnel information sharing, personnel management of all departments, including the modification of basic personnel information, personnel transfer between departments, etc. Through the Data Middle Platform, the superior can centralize the management of the personnel in each department system, which reduces the complexity of the management work.

This paper analyzes the conceptual model of the business system of each department and the existing business in the system, determines which shared data tables exist in the business system of each department, and establishes the data subject model. The process is as follows (Figure 1).

For example, Table 1 below, after reading each business system data table, the data table with the same name is de duplicated. The directory is established according to the data table name set obtained by de duplication. According to the established directory, the data topic model is established, and the data fields in the data table are named uniformly, and the data standard set is established.

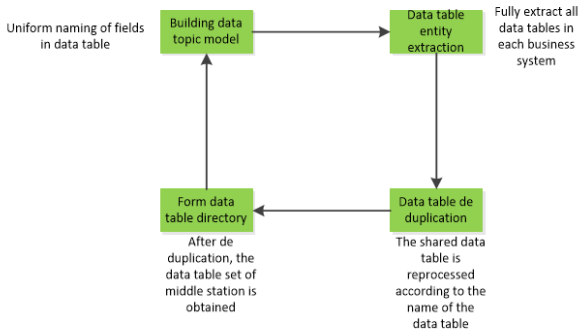

Figure 1: Shared data modeling process

Taking people information management as an example. Mobile phone ID number name, address, ID card number, mobile phone number, gender, Department, user password and so on. The standard set of information data for Party members is shown in Table 1

For example, Table 2 below, the organization activity module is used to display the activities organized by the department. You can view the details of the activities and add the content of the organization activities. Among them, activity information attributes include activity theme, activity content, participants, departments, activity time, location and so on.

For example, Table 3 below, In order to be easy to maintain and update, add data table records the mapping relationship between the various business systems and the data standard set of the Data Middle Platform.

This table is used to explain the mapping relationship between the fields in the original business system and the data standard set in the data central platform system, and records the corresponding relationship between the fields in the data central platform standard set and the fields in the original business system, the actual meaning of the fields, the types of the fields, the positions of the fields, the data sources and the creators of the fields. The administrator in the system can assign authority to the user, so that the user can add new fields in the standard set when needed.

In order to solve the storage problem of a large number of pictures and attachments in the Data Middle Platform system, this paper proposes a method of segment page storage management based on unit. Create a folder for each unit to store pictures and attachments. And when the number of pictures reaches the threshold value, the system will automatically generate a new folder for the unit.The storage principle is as follows.(Figure 2)

A new image storage table is added to store images based on image ID and unit ID. Table 4 below.

Add mapping relationship table between unit and image storage folder. Table 5 below.

There is redundancy between the data from various business systems. To solve the problem of data fusion, the Data Middle Platform fuses the data in the data table with the same functional structure in different systems. The records in each data table have unique identification in different business system data tables. The standard

Table 1: Standard set of people information table

Field name	type	Entity name
id	bigint	Unique identification
units	Varchar	Unit number
account	Varchar	Account
pwd	Varchar	Password
status	Varchar	State
dept	Varchar	Department number
name	Varchar	Name
IDnumber	Varchar	ID
sex	Varchar	Sex
mobile	Varchar	Telephone
Email	Varchar	Email
loginTime	Varchar	Login time
loginIp	Varchar	Login address
isLeader	Varchar	Is he a leader
remark	Varchar	Remarks
birthday	Varchar	birthday
jointime	Varchar	Time of entry
education	Varchar	education
nativeplace	Varchar	Native place

Table 2: Set of standards for organizing activities

Field name	type	Entity name
id	bigint	Unique identification
unitsid	Varchar	Unit number
Tile	Varchar	Activity Theme
publisher	Varchar	Publisher
dept	Varchar	department
time	Varchar	Activity time
publishtime	Varchar	Publish time
reply	Varchar	Number of replies
concerns	Varchar	Number of concerns
content	Varchar	Content
likes	Varchar	Likes
person	Varchar	Participants
place	Varchar	Activity location

Table 3: SYSTEM_ROLE TABLE

Field Name	Type	Entity name
systemId	Varchar	System number
name_new	Varchar	Field names in the standard set
name	Varchar	Field names in business systems
mean	Varchar	Field meaning
type	Varchar	Type
formname	Varchar	Data table name in Data Middle Platform
source	Varchar	Source
person_name	Varchar	Creater
creattime	Varchar	Creattime
tablename	Varchar	Business system data table name

Table 4: Picture storage table

Field Name	Type	Entity name
id	Varchar	Unique identification
picid	Varchar	Picture information number
picurl	Varchar	Picture address
unitsid	Varchar	Unit number
pictable	Varchar	Picture information data table

Table 5: Unit folder storage mapping table

Field Name	Type	Entity name
unitsid	Varchar	Unit number
foldername	Varchar	Folder name
unitsname	Varchar	Unit name

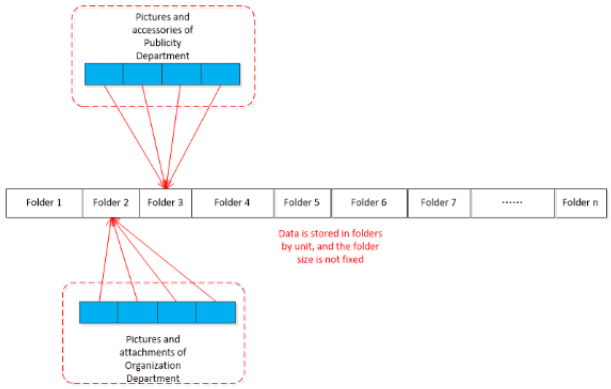

Figure 2: Principle of segment page storage based on unit

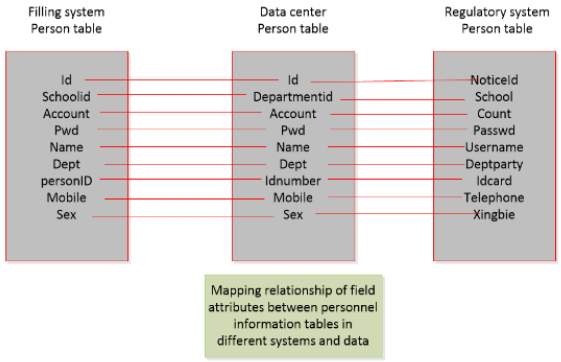

Figure 3: Field mapping

to judge whether two records are the same is to compare whether their unique identification is the same.

OneID refers to a unified data extraction, is a set of ideas and methods to solve the problem of data island, traditional management system in their respective departments to carry out activities, data storage, etc., between each of the data of dispersion, and between different departments of the same definition and implementation of the business entity may, between data cannot be directly associated, this leads to a large number of redundant data. Through the design concept of OneID,a unified entity identification and connection can be built to break the data island in the system and achieve data accommodation among various departments. System such as fill department personnel id number field called PersonID, regulatory system called IDcard, staff id number field in the data in the middle personnel id number field is defined as IDnumber, at the same time will IDnumber logo in the generation of the layers of data into the system, through the id can be linked up each department data domain data, personnel data, accommodation, to ensure that all business processes such as comprehensive and accurate data applications. OneID concept design OneID table structure is shown in the Table 6

Taking the personnel information table as an example, the mapping relationship between the personnel information data fields is reflected through the oneid table.(Figure 3)

4 CONCLUSION

Modern company departments work data and business diversification and scale than in the past, different departments system database information system and service interfaces are mostly independent development, between various departments system caused by such information cannot be interconnected, over time the data island is produced, data and business too scattered result in a large number of redundant construction and waste of resources. In view of this situation, this paper presents the design and implementation of the data middle platform. By middle management system will be the dispersed data island together, unified data standards, at the same time also will be the future continuous data integration to the data in the middle of the join, to meet the future by the concept of continued ascension may meet difficulty, to provide decision-making and support to department management, promote

Table 6: Unit folder storage mapping table

Field Name	Type	Entity name
Onename	Varchar	Data Middle Platform standard field name
Formname	Varchar	Data table name
Name1	Varchar	Field name in filling system
Name2	Varchar	Field name in supervision system
Name3	Varchar	Field name in organization management system
Name4	Varchar	Field name in big data analysis system
Name5	Varchar	Field name in daily assessment scoring system

the standardization of the management work, the intelligence development.

This paper analyzes the demand for data center in the department management system and what shared data exists in each department management system. Data tables of each system are extracted and reconstructed to establish a data topic model. Data standard sets of shared data are respectively established to establish a unified standard. After the completion of data collection, in view of the excessive number of pictures and attachments in the data platform, this paper realizes the storage management of unit-based pictures and attachments, and establishes the picture index based on the unit. Through the design of the OneID table structure, the data within the business system of each department is linked together to achieve the accommodation of data and business.

Current data intelligent middle degree can satisfy the basic needs of each department, China, this paper set up data in the future China can consider to AI and the direction of the algorithm, including using the decision tree and random forest, xgboost logistic regression model building model, and then to optimize parameter tuning method training model, for big data more accurate and in-depth mining, promoting the development of intelligent management.

REFERENCES

[1] Wu H,Shen L,Chen X Z, *et al.* Research on the Application of Data Middle Platform Technology in Integrated Energy Business System[C]. 2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI), 2020:316-319.

[2] Deng J N. Research on Effectiveness of Colleges Ideological and Political Education in the Age of Big Data[J]. Journal of Physics: Conference Series, 2021, 1744(3).

[3] Prakash K U,Malali P,Noronha H. Big data integration for transition from e-learning to smart learning framework[C]. 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC),2016:1-4

[4] Naimi A I,Westreich D J. Big Data: A Revolution That Will Transform How We Live, Work, and Think[J]. Mathematics & Computer Education, 2014, 47(17):181-183.

[5] Al-Bahri B,Noronha H,Pandey J,et al. Evaluate the Role of Big Data in Enhancing Strategic Decision Making for E-governance in E-Oman Portal[C]. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), 2020:914-917.

[6] Agrawal R. Big data: Old wine in new bottle[C]. 2015 IEEE 31st International Conference on Data Engineering, 2015:1597-1599.

[7] Zhou K,Fu C,Yang S. Big data driven smart energy management: From big data to big insights[J]. Renewable Sustainable Energy Reviews, 2016, 56:215-225.

[8] Kim J,Ha H,Chun B G,et al. Collaborative analytics for data silos[C]. 2016 IEEE International Conference on Data Engineering. 2016:218-221.

[9] Han Z and Yu L. A Survey of the BCube Data Middle Platform Network Topology[C]. 2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity),2018:229-231.

[10] Lanka s,Bajracharya s. Incorporating Big Data with IoT in Data Middle Platform Analytics[C]. 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT),2018:240-245.

[11] Moud H I,Hariharan J,Hakim H,Kibert C,et al. Sustainability Assessment of Data Middle Platforms Beyond LEED[C]. 2020 IEEE Green Technologies Conference (GreenTech), 2020, 238-241.

[12] Jia L,Yang H,Huan L,et al. Research on Data Middle Platform Operation and Maintenance Management Based on Big Data[C]. 2020 International Conference on Computer Engineering and Application (ICCEA), 2020:124-127.

[13] Shuping N,Feng W. The Network Architecture Design of Distributed Dual Live Data Middle Platform[C]. 2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), 2019:638-642.

[14] Suresh T,Murugan A. Strategy for Data Middle Platform Optimization : Improve Data Middle Platform capability to meet business opportunities. 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud),2018:184-189.

[15] Yu Y,Cheng G,Qiang X. Data centre transformation: Integrated business model framework of cloud computing oriented data centre[C]. 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016:1848-1855.