Digital platform architecture design based on JavaEE

Hao Li¹ and Heda Yang²
School of Electronic Engineering, Beijing University of Posts and Telecommunication, Beijing, 100876, China
m18515160821@163.com

Abstract. With the comprehensive popularization of electronic technology and the internet, more and more people prefer to use smart phones, tablets and other intelligent terminal devices to read digital information on apps or Web pages, which makes the traditional publishing industry transform to digital service. Based on the platform of JavaEE, this project provides architecture design for the transformation of rail transit publishing house to digital publishing service, integrates the traditional publishing industry resource base, and brings great help to the utilization, optimization and management of knowledge resource base.

1. Project background and significance
The continuous development of science and technology, the comprehensive coverage of the Internet, the perfect combination of new electronic products and network technology have brought many changes to people's life and work. Reading digital content with electronic products can be seen everywhere in our lives. The traditional paper publishing industry has been seriously impacted, and the emerging digital publishing has obtained unprecedented development opportunities. With the rapid development of digital publishing, how to make full use of existing resources and integrate new resources to improve publishing efficiency and quality has become a standard for publishing houses to measure their viability.

With the rapid improvement of computer technology, a variety of service architecture systems have shown up in people's vision. The lack of original functions, such as navigation, index, search and other simple functions of the digital publications service system of rail transit is far from meeting people's needs. Faced with the rapid emergence of massive data, the traditional rail transit digital publication service system is faced with great challenges, such as low efficiency of resource sorting, large amount of resource backlogs, untimely information update and improper writing of data contents caused by the non-standard resource database. So a professional, efficient and intelligent digital system of rail transit services needs to be developed. Its significance not only solves the problem of transition from traditional printing industry to digital publication for publishing house digital service, but also brings a lot of reference significance for other industries to Internet transformation. More importantly, it lays a theoretical foundation for the development, integration of resources and optimization of services in the digital era of big data Internet.

2. Design framework

2.1. System functional architecture
According to users' functional requirements for the digital publishing service system of rail transit and the system's management and archiving of data, we designed the system to be composed of front-end Web pages and back-end servers. Front-end page operation is written by JSP program, and JSP uses
Java language as a scripting language to bring convenience to the programmer. It can receive client requests, and according to the request content dynamically generate HTML, XML or other formats of the Web page[1]. Users access the website to schedule and manage the content of the entire platform. During the operation of the whole system, front page and background constantly interact to achieve data collection, analysis, processing, storage and other functions. System data architecture is shown in figure 1:

The data processing process of this system can be divided into six steps: data collection, data analysis, data processing, data storage, data management and data application. The whole process includes the whole process from receiving data to displaying data.

Data source: the source of the underlying data includes manual online scanning with laser scanner, connection with corresponding business interface for data transmission and import and export of target data. The source of data is the starting point of the whole digital system platform, which includes both various published materials stored by the publishing house and various newly imported information. Data source capacity is a measure of a publisher's access to online and offline resources.

Data analysis: analyse the collected data, and mark them according to machine learning algorithms and big data means according to whether the data are of the same format type, whether the capacity is within the specified range, whether the extracted keywords are the same, and whether the difference coefficient of data comparison is within the specified range. The function of data analysis is to deal with metadata.

Data processing: since data analysis marks each data, the data are classified according to the mark, then the data in different formats are converted into data in a storable format, and they are processed, packaged and compressed, and too large data are split for convenient storage.
Data storage: distributed storage is adopted to achieve load balancing, so as to avoid "hot spot" problems caused by excessive concentration of data, which may lead to server outage and result in data loss or damage[2].

Data management: archive and backup the stored data.

Data application: it is displayed on the front page through the encryption business publishing system platform for users to make statistical analysis, import and export, query and review and other functions. Through the previous processing, the maximum value of data can be brought into play at the application level to realize digital publishing for users in the publishing industry.

2.2. System design scheme

2.2.1. Front-end presentation layer design. Front-end Web page is the entrance of the whole system platform, which should not only meet the needs of users, but also pay attention to the security and authority of information. Here, we put forward the following requirements:

- Different users need to carry out identity authentication when accessing the page login system. If the user information is not matched in the background database, the platform is not allowed to log in. The identity information can be stored in the background database through online registration. After logging in the platform, you can only view the directory within the permission of accessing yourself. Administrators can review resources for data management. Therefore, the system platform needs to support the following functions: user login/registration function, user permission application function, user loan checking and verification function, and administrator to create/delete the resource library function

- Since mass data import and export through Web portal, data transmission efficiency is an issue we must consider. Although HTTP hypertext transfer protocol is simple and time-efficient, it needs to load all data into memory when transferring data, which leads to low efficiency when transferring large files, slow time and excessive system resources. In the transfer of large files FTP file transfer protocol has shown a high advantage, it does not need to transfer all the contents of the load into memory, the transfer of larger files using FTP than HTTP efficiency is higher. Because this system uses ftp file transfer protocol to deal with the problem of big data uploading[3].

- Abide by the principle of page design, to ensure the user guide principle, easy to operate principle, visual balance principle, uniform color matching principle, text readable principle and so on[4]. A good page not only conveys web information to users, but also creates a harmonious, beautiful and pleasant online browsing environment and atmosphere, which improves user experience and guarantees page views for the system platform. Therefore, it is necessary to do a good job in front page design, and the first impression of users is from here.

2.2.2. Background server design. In order to shorten the development time of the system platform and improve the operation efficiency of the system platform, we adopt layered architecture design for the background server to carry out multi-layer decoupling of different system module areas, which not only brings convenience for developers to assign tasks, but also reduces the difficulty of system maintenance and debugging. Background server architecture design as shown in figure 2:
Figure 2. Background server composition diagram.

Database: contains user information tables, user transformation information tables, and store resource data.

Scheduling server: it includes login/registration, account maintenance information, large file splitting, task assignment, front-end page feedback conversion results and other functions.

Conversion server: it includes the functions of accepting conversion requests, performing conversion operations, and feeding back conversion results.

File server: contains the files before the conversion and after the conversion.

In order to prevent data from being deleted, lost or damaged by mistake and improve data security, we adopt distributed data storage and remote data backup, and relational database and non-relational data parallel storage play their respective advantages. This distributed storage architecture is more secure, scalable, and highly available than a single traditional database.

2.2.3. Exception handling mechanism. According to the different abnormal occurrence, the system adopted the corresponding solution. Data transmission is interrupted due to network instability and will continue after the network is restored. User operation error, the background will be in the form of log error information printed and stored in the database; When the server is down, zookeeper coordination service will automatically switch to the backup server to ensure the smooth operation of the platform without terminal[5].

3. Conclusion

This system caters the traditional publishing house for the demand of digital publishing service mode transformation, which has greatly improved the digital publishing quality and efficiency. Under the wave of big data Internet, the traditional publishing houses are bound to transition to the digital industry, which may be a long time, but as long as the technology is constantly improved and user demand is improved, the transition to digital publishing will be successful. According to the plight of the traditional publishing industry, the design structure of this paper combines big data related knowledge and the development architecture based on Spring+SpringMVC+Mybatis, which improves the throughput, diversity demand, security and stability of the system, and has reference significance for the processing of large amount of data load balancing and other fields.

References

[1] Duane, K. (2001) Web Development with Java Server Pages. Manning Publications Co. Greenwich, CT, USA.
[2] Jagadish, H.V., Gehrke, J., Labrinidis, A., et al. (2004) Big data and its technical challenges. Communications of the ACM., 57(7):86-94.
[3] Madhavi V. (2011) Survey of Parallel Data Processing in Context with MapReduce. Comput Sci Inf Technol., 69-80.
[4] Pimentel, V., Nickerson, G. (2012) Communicating and displaying real-time data with WebSocket. IEEE Internet Comput., 16:45-53.

[5] Capitanio, S., Nanni, C., Marini, C., et al. (2015) Heterogeneous response of cardiac sympathetic function to cardiac resynchronization therapy in heart failure documented by 11[C]-hydroxy-ephedrine and PET/CT. Nuclear Medicine & Biology, 42:858–863.