Research on Intelligent Information System of Library under Big Data and Digitization Technology

Chunyu Wang, Zhengyu Sha*

Library of Wuhan University of Technology Wuhan University of Technology
Wuhan, Hubei, China

*Corresponding author: zysha@whut.edu.cn

Abstract. In the big data environment, we develop personalized information of college libraries based on big data from three aspects: the overall architecture of the system model, the functional model of the system, and the design of system interface modules according to the design principles and requirements of the personalized information service system of the university library Service system design. In terms of the functional design of the platform, the service platform is divided into four levels: accurate identification of user needs based on big data, personalized customized services based on artificial intelligence, academic research and discussion space based on integrated media, and fine-grained subject resource aggregation based on knowledge. On this basis, a centralized model of individualized services of university libraries including internal and external personnel, information resources, technology, services, processes, platforms, and environment has been constructed.

Keywords: Big data technology; colleges and universities; library information system.

1. Introduction

The rapid development of information technology characterized by computer technology and network technology provides automated, networked, and digital means for the library's value positioning, service mode, organizational restructuring, and many other reform issues. Especially in the new user information environment constructed by the Internet, the contradiction between the initiative and selectivity of users to obtain information resources has become increasingly prominent. In this kind of information environment, how college libraries can adapt to this change, and it is particularly urgent to establish a user-centered information service model [1]. Only the information system selected and used by users has meaning and value. Therefore, only by truly realizing the transformation from the traditional library information service model centered on the collection of documents to the user-centered modern information service model, can we better exert the value of the collection; on the other hand, information technology also implements the user-centered approach. The information service model provides the corresponding technical conditions. At the same time, the customer relationship management theory widely adopted by the business community provides a methodological basis for exploring the establishment of a user-centered information service innovation model.
2. Analysis of the principles of platform construction

2.1. The principle of deep openness to teachers and students centered on intelligent service users
Traditional subject service platforms are often built around libraries, ignoring the experience of platform users. Based on the “Internet+” thinking, the intelligent subject service platform focuses on intelligent service users, and uses target embedding, function embedding, process embedding, system embedding, space-time embedding, ability embedding, and ability embedding based on resource service to people [2]. A variety of ways such as emotional embedding and collaborative embedding are embedded in each stage of teacher and student learning.

2.2. The principle of intelligent resource retrieval based on the convergence of subject resources
To evaluate the quality of subject services, first, we need to measure the ability of the library to gather resources. The library's intelligent subject service platform should be built on the basis of gathering abundant resources, and reasonably use cloud storage and cloud computing technologies to realize intelligent classification and clustering of unknown network resources, assist users in obtaining the effective resources they need, and provide users with a unified and standardized search portal, through rigorous data merging, deduplication and sorting, provides users with logical and meaningful search results, builds a subject-oriented and user-oriented resource cloud platform, reveals and obtains various types Distributed and heterogeneous resources, realize unified browsing and retrieval and target information acquisition, and promote users to realize platform intelligence needs in the case of fuzzy information.

2.3. The principle of comprehensive application of mobile terminals with MOOC as the main line
Traditional library first subject service platforms mostly organize resources by professional classification such as Chinese Library Classification. For user groups, what they need most is the resources most closely related to their own research fields. Therefore, the intelligent subject service platform needs to use the school subject construction system as a reference to organize resources, and use MOOC as the main line to gather resources. This is more convenient for the research and learning of the student group, and it is also easier to recommend resources for the courses taught by excellent teachers [3]. In addition, as the proportion of mobile smart terminal devices such as smartphones and Pads owned by students and teachers is increasing, their learning methods are also biased towards mobile and fragmented learning. The full application of mobile terminals can enable users to carry out anytime, anywhere learning, truly realize ubiquitous learning that allows users to be free from any restrictions.

3. System design requirements
It was decided that the design of a personalized information service system for university libraries based on big data should meet many requirements: First, the personalized information system should have a good user interface. As the interface between the user and the personalized service system, the user interface should be simple, clear, clear, and beautiful, with help prompts designed to meet the real needs of users at different levels, express the functions of the system, and help users solve problems in use. Second, high-efficiency and personalized information services should be provided [4]. Users use the personalized information service system of university libraries based on big data to obtain and filter information, obtain the best information in the shortest time, and improve service efficiency. Finally, provide personalized information services with different service depths. For user groups with high service depth, high-quality personalized information services are required, which can be customized by referring to user opinions during development; for user groups with low service depth, templates can be set up in advance during development. The selection of information can be prompted through the interface to provide flexible and personalized information services for high and low-level user groups to facilitate convenient use by user groups at all levels.
4. Library personalized service system architecture
The library is based on the application of RSS, Blog, Wiki, Tag, IM, and Facebook in the library as its main form. Its system architecture can fully demonstrate the overall pattern and planning of library information services, and is to construct and deeply understand books. The basis of the library’s personalized service model. The library's personalized service system architecture is the overall structure and framework of the library's personalized service. It consists of a four-layer structure of technical foundation, information resource construction, personalized service platform, and demand-based personalized service application (see Figure 1).

![Figure 1. Library personalized service system architecture.](image)

4.1. The basic layer of library technology
This is the technical basic environment of the library architecture, which consists of network facilities, technical support, and management mechanisms. The network facilities are composed of software and hardware of library network technology and computer communication technology. Technical support mainly includes platform management technology, network security technology, information generation, processing and storage technology, multimedia database technology, data mining technology, information visualization technology, metadata technology and related standards and specifications [5]. The management mechanism is the management technology and means to ensure the stable and effective operation of the system, including the policy orientation of platform construction, capital investment, and regulatory evaluation.
4.2. Library information resource construction layer
It is composed of two-way resource integration mode, resource construction six steps and management feedback in the library environment. The two-way resource integration mode of the library, that is, the user participates in the resource construction mode from the bottom up and the library top-down resource construction mode. The six steps of resource construction are information resource planning, production, selection, collection, organization, and storage. Management feedback is to supervise and control the resource construction model and all links.

4.3. Library Personalized Service Platform
The platform is to provide a shared space for university users, mainly composed of internal and external personnel (librarians, users, and experts), technology, resources, services and processes and other core factors.

4.4. Demand-based library personalized service application layer
This level is a library personalized portal based on the above core elements. It is supported by library tools and user demand mining technology to ensure the effective operation of the library to provide users with personalized information push, personalized information customization and other services. It is the top layer of the library structure and the overall display of other hierarchical structure functions.

5. Service platform construction technology

5.1. Functions of Subject Service Platform of University Library
In terms of the service form of traditional university libraries, physical libraries or network libraries are the mainstay. In the era of big data, important changes have taken place in library service methods, service channels, and service models, and a new type of data library has begun to form. University libraries can easily collect big data, and then process them on this basis, and finally perform reasonable intelligent services. The service function is shown in Figure 2.

5.1.1. One-stop service for data resources. The various data resources collected by the library are massive, including various documents, academic conferences, scientific research results, etc.; there are also various records of students and teachers; it also includes library services and announcements.
However, because these data may come from different places, we can make full use of the advantages of the relevant libraries of the library, and then organize the analysis and connection of various data sources to realize the fine operation of big data. Because these resources have shortcomings such as huge information, multiple types, and no order, we need to establish a unified standard to meet this requirement. We provide a variety of resources for seamless interconnection, and provide multiple effective services for data. In the torrent of data, we must converge effective resources, use pre-indexing methods to form a highly concentrated central knowledge system library, provide users with simple and efficient search and access services, and establish a one-stop service platform.

5.1.2. Information Visualization Service. Big data contains a lot of important information, and people look forward to in-depth analysis in order to better discover knowledge from the data. Information visualization is a key service concept and technical means for complex processing of library big data, and it can give users a relatively uncomplicated knowledge environment [6]. It is possible to visualize the topic from the perspective of finding their relationship from a variety of data results, and then realize the complex visualization of multiple relationships. It is also possible to combine visualization and measurement methods and other scientific methods, and then generate maps with scientific attributes to explain related field structures and development trends, promote information acquisition, make knowledge structures more obvious, and express the relationship between multiple disciplines. In the user scenario, the complex and invisible semantic relationship data is displayed in the way of information visualization. Let the people who use it organize the information effectively and then fully analyze and utilize the value hidden in the big data information resources.

5.1.3. Subject knowledge service. As a knowledge service platform, the library improves the value density of data through processing and sorting of collected data and data modeling. Based on data, the library develops knowledge services through data-based value-added services to meet users' academic knowledge needs. For example, based on disciplines, classify the information behavior data of users in different disciplines; analyze the relevant documents searched, browsed, and downloaded by users, add relevant tags, and then summarize the user's hobbies in this related field; use big data-related technology for prediction and analysis. Build a clickstream book borrowing data warehouse based on big data logs, conduct relationship analysis to discover the network knowledge system between them, and create related dimensional models; use big data subject trend analysis to improve book services through orderly processing of big data quality.

5.1.4. Smart Service. The increasing demand of modern readers for knowledge makes the knowledge service of the library progress in the direction of individualization and intelligence. Personalized service analysis and based on the specific needs of users, and then proactively provide readers with relevant information that they may be concerned about. Intelligent service is the sublimation and core of knowledge service. The library stores a large amount of data generated by user information behaviors, such as logs generated by users querying bibliography, circulation log data generated by users borrowing and returning books, log data generated by users searching, browsing, and downloading electronic resources, traffic data generated by user access, and various in addition to recording the personal information of readers, these data also hide a lot of important information. Therefore, the relevant analysis technology of big data can be used to meet the needs of users for intelligent services.

5.2. System interface module design

5.2.1. Login module design. This module is mainly used in the process of collecting personal information of new user groups. New users can register for an account, fill in personal information, and submit it to the system for login, as shown in Figure 3.
5.2.2. **User customized service interface design.** This module is mainly to provide registered users with the required system service interface, collection of resources and processing of search results, etc., according to different levels of service objects, different login page settings, which is shown in Figure 3 in the system User interface 1 and user interface 2 shown, the system can operate according to whether personalized service information is customized, through mobile phone number, email address or security question identity verification. If personalized service information needs to be customized, enter the user interface 2, and vice versa enter the user interface 1, in the interface, the content of the information service provided by the system can be appropriately modified and managed. At the same time, by comparing the service content of user interface 1 and interface 2, user interface 2 has latest announcements and "my resource library" service content, and users can learn the latest cutting-edge information and can obtain the latest information resources for themselves. Add, delete, store, backup, import and export, etc., so that "My Resource Library" can reduce junk information and store the latest and most important information resources.

5.2.3. **Evaluation Module.** The implementation and evaluation sub-framework describe the content that needs to be focused on during the implementation and evaluation of data governance, involving the environment, steps, and effect evaluation required to implement data governance. It mainly includes three aspects: implementation methods, maturity evaluation, and auditing (see Figure 4).
5.2.4. Personalized reference consultation module. The online survey found that the traditional reference consulting methods of university libraries mainly focus on E-mail, telephone and BBS, and face-to-face communication between users and librarians. In the library environment, university libraries can provide users with reference consulting services in a variety of ways through technical services such as blogs, Weibo, wikis, IM, and SNS represented by Facebook, and information aggregation, tags, and browser plug-ins and other technologies are more embedding the above technologies to provide push and organize information resource services. In the questionnaire survey, university users and librarians have a low level of knowledge and application of library reference consultation, but they are full of expectations for library personalized reference consultation services: 91.3% and 68.9% respectively. Users expect the library to provide online communication forum services and consulting blog services. To this end, we have constructed a demand-based library personalized reference consulting service model (see Figure 5).
5.2.5. Actively collect and filter information module. The information filtering method based on the keyword vector is a similar filtering method. The idea is to compare the similarity between the keywords that characterize the information unit and the standard keywords used for filtering. Only when the similarity reaches a certain degree, the information unit is retained. Establishing the similarity of keywords is the core of this filtering method. Both the keywords of the information unit and the standard keywords can be regarded as vectors, and the similarity is the cosine value of the angle between the two vectors. If \( W_j \) is the keyword weight vector of information unit \( j \), and \( P \) is the standard keyword weight vector. The similarity \( R_j \) between the two is expressed as:

\[
R_j = \cos(W_j, P) = \sum_{k=1}^{k} \frac{W_{j_k}}{|P||W_j|}
\]

The information filtering method based on the document set is a comparison method based on the training set. The Information Agent analyzes the information documents provided to users by the previous service system, determines the information documents that users like, uses it as a training set, and classifies the information documents newly acquired by the system based on this, and divides them into information documents that the user likes and users. Information documents that users don't like can be removed by removing information documents that users don't like to achieve the purpose of information filtering.

6. Conclusion
The advent of big data will inevitably bring significant development opportunities to the library's subject services. Only by making good use of relevant technologies can university libraries provide better convenience for users with professional backgrounds. Subject services in the context of big data should
be developed around user data. Subject librarians can use cluster analysis of user data to refine user needs, collect and organize resources and feed them back to users, thereby completing targeted services. In order to realize the subject service function under big data, it is imperative to build a corresponding subject service platform. This article discusses the specific construction technology of the subject service platform from the three aspects of platform function, platform architecture and big data technology. With the help of these technologies, the library can make full use of users’ data resources, discover potential subject needs and hidden user networks, so that communication is no longer limited to the straight-line model of subject librarians and users, but users and users, users, and users. The network mode of subject librarians promotes the comprehensive and perfection of the service functions of university libraries.

References
[1] Horova, S. Informatization and society: possible difficulties and future prospects. Наукові праці Національної бібліотеки України імені В.І Вернадського, vol. 5, pp. 41-57, August 2020.
[2] McFadden, C., Alpysbaeva, N., & AImgazinov, K. INFORMATIZATION OF ARCHIVAL BUSINESS IN KAZAKHSTAN (foreign and domestic experience). Вестник КазНУ. Серия историческая, vol. 93, pp. 22-29, February 2019.
[3] Tsarapkina, J. M., Petrova, M. M., Mironov, A. G., Morozova, I. M. & Shustova, O. B. Robotics as a basis for informatization of education in a children's health camp. Amazonia Investiga, vol. 8, pp. 115-123, February 2019.
[4] Bazhenova, S. A. Approaches to improving the training of teachers working under the International Baccalaureate programs in the field of education informatization. RUDN Journal of Informatization in Education, vol. 17, pp. 123-133, February 2020.
[5] Levchenko, I. V. & Sadykova, A. R. Approaches to solving the problem of search of scenarios of lessons on informatics for basic school in the Moscow E-School library. RUDN Journal of Informatization in Education, vol. 16, pp. 231-242, March 2019.
[6] Shvartsman, M. E. & Kryzhanovskaya, O. N. Digital Research Methods: the New Vector in Library Informatization. Bibliotekovedenie [Russian Journal of Library Science], vol. 68, pp. 103-111, January 2019.