Design of Digital Library Data Search Engine Based on Cloud Computing in Big Data Era

Baoyan Tang and Bingfu Hu*
Weifang Engineering Vocational College, Qingzhou, Shandong, China

*Corresponding author e-mail: hubingfu@wfec.edu.cn

Abstract. In the big data environment, people's acquisition of knowledge is no longer limited to the contents of books, but more data information from the Internet. Intelligent search engine implements information retrieval based on knowledge level, with strong natural language understanding and knowledge processing ability, showing good personalized information service characteristics. He theory of information ecology uses the theory, viewpoint and method of ecology to study the constituent factors of information ecosystem. Excess information and meeting the needs of readers have become the main contradictions in library services. As an architecture model suitable for library application, cloud computing can integrate scattered internet information data, realize the intensification of digital libraries, and provide a new solution for building and sharing digital resources and intelligent search engines. This paper discusses the construction mode of digital library service platform based on cloud computing technology, and analyzes the application and design method of intelligent search engine of cloud service digital library under the background of big data.

Keywords: Search Engine, Big Data, Cloud Computing, Digital Library

1. Introduction
Information resources are characterized by high digitalization, networking and multimedia. However, faced with the vast amount of information, people suddenly find it difficult to find the information they really need [1]. In the big data environment, people's acquisition of knowledge is no longer limited to the content of books, but to obtain more data from the Internet [2]. Big data technology is effectively applied to digital library information services. It retrieves useful information resources from complex and multi-structured massive data, and provides the obtained information resources to users in a visual way. It is a digital library. As a social information service organization, the library provides the most reasonable structure, the most comprehensive disclosure, the most scientific organization and the most continuous category of information resources by comparison [3]. The university library website has become the bearing place of university knowledge, and it has become a broad platform for university libraries to provide users with information resources services and interactive exchanges [4]. How to solve the problem of information imbalance and non-ecological problems, many scholars have used the theory and methods of natural ecosystems to explain and study...
and have achieved certain achievements, which has also led to research topics on information ecosystems [5].

How to effectively manage the massive information resources of the digital library and further transform it into the actual needs of users has become a realistic issue of concern to the industry [6]. People have changed the previous single way of obtaining information knowledge from books, and obtained richer, more accurate, more timely and more suitable information from the network [7]. As an architecture model suitable for library applications, cloud computing can integrate scattered internet information data, realize the intensification of digital libraries, and provide new solutions for the joint construction and sharing of digital resources [8]. Intelligent search engine, as the third-generation search engine produced in the development of contemporary IT technology, has good natural language understanding and knowledge processing capabilities, and has strong advantages in intelligence and humanization in the information service process. Intelligent search engines are currently It is widely used in many kinds of service systems [9]. In order to meet the different information needs of various users of university libraries, libraries should understand the characteristics of users in a timely manner and analyze the information needs of users. "Cloud library" can improve the utilization rate of information resources, realize resource sharing, and improve the satisfaction of users. The theory of information ecology is based on the research of information science and ecology [10]. This paper discusses the construction idea of digital library service system based on cloud computing, and analyzes the application and design method of intelligent search engine of digital library based on cloud computing in big data environment.

2. Characteristics of Library Data Environment in Big Data Era

Digital library is a collection of massive digital resources which has nothing to do with the platform. The information service of digital library can be realized between databases and workstations with good interoperability between services. At present, digital library is actively exploring information services with deep semantic interoperability. In the big data environment, user push service with personalized customization as the core has become the main direction of digital library service mode reform. In order to meet readers' digital reading needs and improve users' satisfaction, the library will share service data and user data resources with cloud service providers, e-commerce operators, third-party value-added service providers, communication service providers, etc. in the way of data sharing on big data platform. Digital library has powerful information dissemination and publishing functions. Compared with traditional library, digital library has significant differences in service, which abandons the passive service of traditional library and introduces active service. Through the mining and analysis of big data, the library data information needed by users can be retrieved, and targeted library information services can be provided for users.

Under the big data environment, the data mining ability of digital library is put forward more stringent standards. It uses the theory, viewpoints and methods of ecology to study the components of the information ecosystem. Through data standardization work, the overall data planning is more solid, so that the overall data planning results can play a guiding role in the construction of integrated information systems. This is what we call information resource planning. The information organization and storage structure of digital library will become more and more complex, including not only traditional structured single media information, but also semi-structured and unstructured media information [11]. The digital library we discuss is essentially an artificial system based on computers and various information technologies, supported by information resources and with people as the core. The ecosystem service evaluation system evaluates a knowledge service according to the feedback of knowledge users and the solution of users' problems and the service process, including the evaluation of information resources, the evaluation of information ecological subjects, and the evaluation of service strategies. Since the amount of information in digital library itself is a large data set, it will be a larger data set if the information of many digital libraries is integrated together. Figure 1 shows the construction of digital library resources under the information environment.
Figure 1. The construction of digital library resources under the information environment

Information resource planning can effectively solve the contradiction between the limited life cycle of information resources and long-term needs. However, among different government agencies, service providers and enterprise big data platforms, there may be some problems, such as inconsistent data storage and management standards, lack of horizontal and vertical communication of data between different big data platforms, information blind spots in platform data management and search, and lack of big data management and integration tools in libraries. Some important data related to the effectiveness of users' personalized service are stored in the large databases of the government, third-party value-added service providers and operators, and have a strong monopoly with the rapid growth of the quantity and value of the data they own [12]. The arrival of the era of knowledge economy has brought the society to a knowledge-based society. Knowledge learning has become the lifelong need of the members of the society, and knowledge innovation has become the main theme of the knowledge-based society. Through the design of basic functions such as data mining and data retrieval, we can provide users with the needs of data collection and storage, ensure the safety and accuracy of data, and serve users better.

3. The Influence of Cloud Computing on Library Information Service

In the previous service mode of the library, effective information is scarce, which makes it difficult for people to find it. Therefore, the library should introduce new related information technology, so that people can obtain information that can better meet users' requirements in the library. Its powerful data storage environment and network services can find solutions to the problems in the operation and service of the library. The library takes measures such as training and incentive to strengthen the management of information and personnel, so as to further upgrade the management mechanism, so as to improve the information system of Digital Library and achieve the goal. From the perspective of users, users are not required to participate in information services when they query information in traditional libraries, while users are required to participate in information services when they query information in Digital Libraries in order to make library information match the effective information needed by users as soon as possible. Information is the fact that people are given a certain meaning and interconnection after systematically organizing, organizing and analyzing data through human cognitive ability. Higher than knowledge is wisdom. Wisdom condenses past knowledge and experience, showing people's insight and forward-looking power toward the development of things and things.

Through microblog, forum and other network media, the sharing platform enables librarians to communicate with readers in real time, express opinions, put forward suggestions and improve services. Such interaction can be between librarians and users, experts and librarians, experts and users,
users or librarians, users and experts. Library can integrate and apply the information resources stored in the "cloud" by taking advantage of the massive storage advantages of cloud computing. Through the Internet, cloud computing can integrate discrete, diverse and heterogeneous information into an orderly, massive and standardized perfect system with powerful computing power [13]. If the users of digital library directly participate in the information service, the librarians can make clear the needs of users, provide better information resources for users, meet the needs of users, and supplement the integrity of digital consulting service [14]. Users can log in to the home page of the website and connect to the cloud service platform to get the services they need.

With the development of the times, there are more and more information resources in the library. These isolated information resources also bring great inconvenience for users to find and obtain the information they need. Table 1 shows the accuracy statistics of the data set in different dimensions. The accuracy numbers of various functional sizes are shown in Figure 2.

**Table 1.** Results of informatization experiments based on information ecology

| Feature dimension | 20   | 40   | 60   | 80   | 100  |
|-------------------|------|------|------|------|------|
| Accuracy (%)      | 78.18| 75.37| 79.58| 82.40| 71.90|

**Figure 2.** The changing trend of informatization construction pattern under different characteristic dimensions

Cloud storage services and software services of cloud computing can integrate various databases and scattered information data, which can complement resources and eliminate "information islands". This layer provides the resources needed for application running, management, monitoring and maintenance. Users evaluate customized information, and digital library adjusts the content provided by users' evaluation, so as to provide readers with the information they really need. In this era of information resource sharing, cloud computing has opened up a new field of information resource sharing for libraries. Constructing information sharing space between different libraries to form a "cloud library", and participating libraries can obtain information resources of other libraries in the "cloud library". By sharing resources among libraries, readers can get more information and enjoy better services.
4. Personalized Service of Library Based on Intelligent Search Engine
At present, the function of library in pushing and recommending personalized service information has been relatively perfect, but the function of knowledge decision-making is still relatively lacking, so it can effectively reflect the role of intelligent search engine in knowledge mining and data discovery. Users of remote information outside the library cannot come to the library in person, but with the help of the network information service platform provided by the university digital library, they can access and log on the homepage of the university library remotely to realize the information demand. At present, when using intelligent search technology, libraries have realized the integration of search results of various search engines to a certain extent, but their integration function is not high enough, so the search results are not ideal in many cases. Therefore, the application of intelligent search engine in library should be actively developed towards integration and intelligence [15]. Internet is also an important way for us to obtain information. Under the support of cloud computing technology, not only all information resources in the library should be integrated, but also network information should be integrated and stored in the cloud.

Intelligent search engine has better ability to deal with and judge library users at the knowledge level, but the ability to analyze and understand knowledge at the ontology level is relatively weak. At the same time, collect metadata from each node cloud database and update the form of data information. The ecological information management subsystem will make corresponding decisions according to the evaluation results of the ecological service evaluation subsystem, including some adjustment decisions, optimization decisions, etc., and transfer the decision information to the ecological knowledge management database [9]. Since data, information and knowledge are the basic raw materials on which library services depend, the comparative analysis of information services and knowledge services should start with the relationship between data, information and knowledge. At present, when libraries use intelligent search engines, the retrieval methods and uses are not rich enough, and there are still many links to be improved in push service and personalized recommendation. Therefore, when applying intelligent search engines, libraries can apply more special features or multiple retrieval methods.

5. Conclusions
Based on the increasingly modern digital library, the demand for personalized service is becoming more and more urgent. Libraries need to make full use of modern information technologies such as intelligent search engines to innovate service functions, and accurately obtain the personalized differences of computer users in retrieval habits, use preferences, learning background and so on. The ecological information management subsystem will make corresponding decisions according to the evaluation results of the ecological service evaluation subsystem, including some adjustment decisions, optimization decisions, etc., and transfer the decision information to the ecological knowledge management database [9]. Since data, information and knowledge are the basic raw materials on which library services depend, the comparative analysis of information services and knowledge services should start with the relationship between data, information and knowledge. Adhering to the concepts of safety, efficiency, intelligence and extensibility in the design and implementation of search engines can ensure the comprehensive, efficient, accurate and economical process of big data search, mining and information discovery, and provide scientific, comprehensive, economical and reliable data support for library big data management and personalized service for readers.

References
[1] Yu Qing. Digital library construction under the background of big data. Management Observation, vol. 611, no. 12, pp. 55-56+59, 2016.
[2] Cheng Luode. Research on Digital Library Information Security Strategy in Big Data Environment. Journal of Library Science, vol. 298, no. 1, pp. 79-84, 2020.
[3] Han Xiangrong. Path analysis of the construction of a new digital library under the background of the public cultural service system. A Comparative Study of Cultural Innovation, vol. 118,
no. 10, pp. 185-186, 2020.

[4] Huang Wenjuan. Application of search engine based on artificial intelligence in digital books. Information Technology, vol. 43, no. 11, pp. 93-96+102, 2019.

[5] Duan Ruibo. Analysis of Big Data Cloud Computing Technology and Its Application. Digital Technology and Application, vol. 342, no. 12, pp. 237-238, 2018.

[6] Yang Fang, Shu Xiaochun. Research on the construction of cloud service digital library. Journal of Agricultural Library and Information Science, vol. 28, no. 9, pp. 24-27, 2016.

[7] Lu Lu, Gao Yihong. Application of Intelligent Search Engine in Personalized Information Service of Digital Library. Information Management of Sci-tech Literature, vol. 128, no. 4, pp. 34-37, 2018.

[8] Xie Feifan, Wang Xiaoyan. Security problems and countermeasures of digital library virtualization based on cloud computing. Information recording materials, vol. 21, no. 1, pp. 170-172, 2020.

[9] Huang Yiming. Digital library information resource integration and service model based on cloud computing. Journal of Library Science, vol. 41, no. 4, pp. 122-126, 2019.

[10] Gong Quansheng. Analysis of the Construction of University cloud service digital library. Information Recording Materials, vol. 20, no. 1, pp. 142-143, 2019.

[11] Li Jin. Design of an intelligent question answering system for digital libraries based on cloud computing. Manufacturing Automation, vol. 41, no. 8, pp. 136-140, 2019.

[12] Chen Ping. The Application Path of New Media in the Construction of University Digital Library. Digital Communication World, vol. 179, no. 11, pp. 136-137, 2019.

[13] Kong Ning, Zhang Changjie. Usability evaluation of smartphone search engines. Digital Library Forum, vol. 10, no. 10, pp. 68-72, 2015.

[14] Zhao Yuanguang, Sun Xiaofeng. Research on Digital Library Service Based on Cloud Computing. Science and Technology Information Development and Economy, vol. 1, no. 11, pp. 59-62, 2016.

[15] Xie Feifan. Research on the Integration of Digital Library Information Resources Based on Cloud Computing. Information Recording Materials, vol. 20, no. 9, pp. 185-187, 2019.