

Current Situation and Development of University Libraries' Self-Built Resources in the Big Data Environment—A Case Study of Self-Built Database of Public University Libraries in Hubei Province

Zhengyu Sha¹

¹Library of Wuhan University of Technology, Wuhan, Hubei 430070, China

*Corresponding author. Email:4261731@qq.com

ABSTRACT

The self-built resources of university libraries are an important part of the digital resource system of libraries. It is a long-term and necessary work to study sustainable development. This work investigated the self-built resources of 36 public undergraduate colleges and universities in Hubei province by means of website visits and network interviews, finding that there are some problems in current construction, such as outdated theme, backward technology and low utilization rate. It put forward transformation strategies such as introducing new artificial intelligence technology, reorganizing library big data, innovating library services, positioning the development direction of self-built resources, and collaboratively constructing university big data, so as to provide reference for the transformation and construction of self-built resources to adapt to intelligent technology.

Keywords: University library, Database, Self-built resources, Big data, Artificial intelligence

1. INTRODUCTION

In the 1990s, the development of Internet technology has promoted the rapid growth of global network information resources, followed by the promotion of open access to academic publishing in the publishing and academic circles, which created more favorable conditions for free access to academic resources. Under this kind of resource environment, the domestic university library relies on the university's superiority discipline, special collection, and characteristic humanities, collects various network related resources extensively, and forms the characteristic self-built database, so as to provide service for the university and the society. By 2016, the number of self-built databases of 39 "985" university libraries in China had reached 382[1], reaching the peak of self-built database. Over the years, although the number of self-built resources in colleges and universities has accumulated a certain scale, the construction model and service function have gradually lagged behind the development of technology, which is far from meeting the needs of users, and a large number of self-built databases are constantly being eliminated. By 2019, the author has counted the self-built database of "985" university library again, and the number is less than 190.

In the past 20 years, the self-built database construction of university library has experienced the process of exploration, development, maturity and gradual decay [2]. The author made a statistical analysis of the self-built

database research literature in colleges and universities, and finds that the research peak is between 2008 and 2014, and then decreases year by year [3]. The research is basically consistent with the actual construction of a self-built database. By analyzing the literature content, it is found that the previous research results mainly focus on the construction theme, resource types, construction standards, investigation and analysis, and evaluation system of self-built resources, while the later research focuses on theoretical research and digital humanistic construction of characteristic collections. There is little literature on the transformation and construction of library self-built database under the new technology environment.

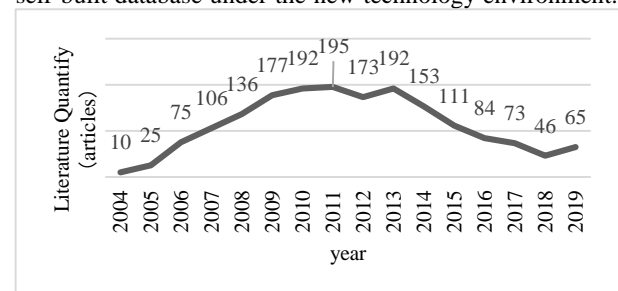


Figure 1. Statistics on the number of resources research documents of Chinese CNKI university library self-built resources

Big data can realize the intelligent processing and application of data, information and knowledge units through algorithms, which provides a new technical means for knowledge management and service in modern

libraries. In big data environment, the old self-built database cannot meet the needs of development in platform construction mode, data acquisition scope, resource mining and user service technology. This work discussed the application prospect of various artificial intelligence technologies in the construction of self-built library resources, the scope of big data of the library and the transformation strategy of service innovation development. As a necessary supplement to the digital resources of university libraries, it is imperative to build a self-built database on the basis of new technology.

2. RESEARCH ON THE CURRENT SITUATION OF SELF-BUILT DATABASE CONSTRUCTION

This work selected 36 libraries of public universities in Hubei province published by the Ministry of Education, and conducted network survey. Statistics show that among the libraries surveyed, 24 libraries have self-built resources and 120 self-built databases (as shown in Table 1). One of the institutions is not accessible and no statistics are available. 18 libraries are setting up a column on the library homepage to display their own resources. Among them, Wuhan University and Hubei University of Economics have the largest number of self-built databases, with 10 each.

Table 1 Statistics on the number of self-built databases in public universities in Hubei province

Colleges and Universities	Number
Wuhan University	10
Huazhong University of Science and Technology	4
Central China Normal University	4
Wuhan University of Technology	3
China University of Geosciences (Wuhan)	1
Zhongnan University of Economics and Law	8
Huazhong Agricultural University	7
South-Central Minzu University	7
Hubei University	1
Wuhan University of Science and Technology	5
China Three Gorges University	8
Yangtze University	0
Jiangnan University	0
Hubei University of Technology	5
Wuhan University of Engineering	7
Wuhan Textile University	5
Hubei University of Chinese Medicine	0
Wuhan Polytechnic University	5
Hubei University of Automotive Technology	4
Hubei University of Education	0
Hubei University of Medicine	5
Hubei University of Economics	10
Wuhan Institute of Physical Education	8
Hubei Normal University	0
Hubei University of Police	--
Hubei Institute of Fine Arts	1
Wuhan Conservatory of Music	5
Hubei University of Arts and Science	0
Hubei Engineering University	1
Hubei University of Science and Technology	1
Huanggang Normal College	0
Hubei Polytechnic University	0
Hubei Minzu University	5
Jingchu University of Technology	0
Wuhan Business University	0
Hanjiang Normal University	0
Total	120

2.1. Resource type

According to the data types of self-built database, it is divided into three resource types: multimedia, secondary literature and full-text database. The number of databases of each type is shown in Figure 2.

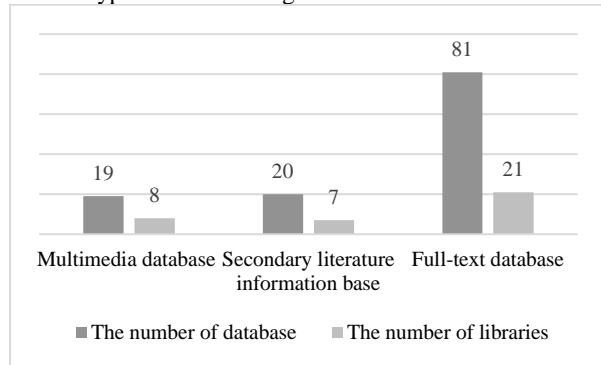


Figure 2. Statistics on the number of databases of various resource types and the number of distributed universities

Multimedia database includes PPT courseware, pictures, audio and video resources. The secondary document information base contains catalogue, item, index and abstract. The full-text database includes an academic thesis, institutional repository, and subject characteristic database. The results show that the self-built resource is an important supplement to the library resource system, and the full-text database is the key content of construction.

2.2. Resource themes

According to the survey, the self-built database has a variety of topics, which can be summarized into seven categories after content analysis: (1) dissertation database, collection of master's thesis, bachelor's thesis and doctor's thesis; (2) institutional repository, including institutional knowledge base and academic achievement database; (3) collection characteristic database, including special collection, library training resources, library publications and book-accompanying CD-ROMs; (4) thematic database, including subject information portal and subject characteristic database; (5) teaching support database, including examination materials, course reference and fine courses; (6) open access database, including free academic resources and free network resources; (7) other types include teacher resource base and regional cultural characteristic resources. The number of databases of each type is shown in Figure 3. The results show that subject resources are the focus of university library construction, and the largest number of libraries using open-access resources, reflecting the importance universities attaches to open access resource utilization.

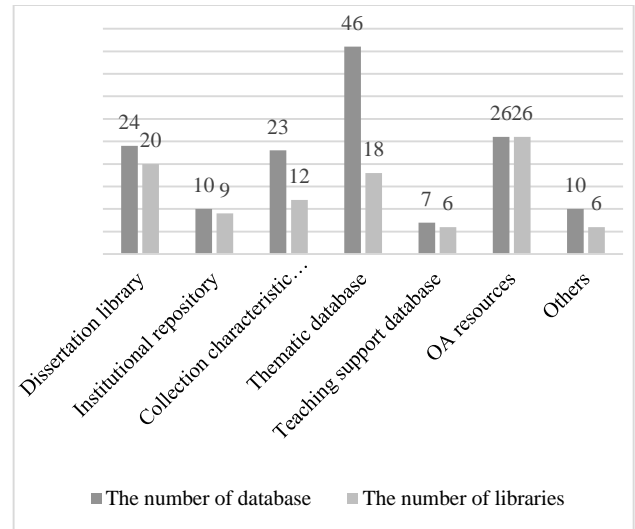


Figure 3. Statistics on the number of databases of each subject type and the number of distributed universities

2.3. Platform function

It is found that some database only provides resource browsing function, some database only provides simple retrieval, very few databases can provide advanced retrieval, and no database has cross-database retrieval function. The survey also shows that some databases once had RSS push function, but then stopped for security reasons. A few databases also have simple statistical functions, such as visitor statistics and high-frequency search term statistics.

2.4. Access channels

According to the survey, only 13 of the 120 databases provide open access channels, while the other databases are restricted to access within the IP address range of the campus or could only be accessed through user authentication. All databases are accessible through mobile browsers. At present, no university library in Hubei province offers mobile client service. The WeChat official number of each school also does not provide self-built resource access interface.

3. ANALYSIS OF THE CURRENT SITUATION OF SELF-BUILT DATABASE CONSTRUCTION

3.1. Results and experience of database development

Over the years, colleges and universities in Hubei Province have built a database with various data types and rich

subject content, accumulating a large amount of data. As a powerful supplement to the digital resources of colleges and universities, which meets part of the user's demand for resources. Taking the library of Wuhan University of Technology as an example, the new material composite technology information portal has accumulated more than 10 million visits [4]. It provides teachers and students with subject news, conferences, papers, institutional information and other resources services. Librarians have accumulated rich experience in building their own databases. For the formulation and adjustment of resource collection strategy, database maintenance and management, data expression, and user need analysis, it has the ability to find problems, ask problems, and solve problems.

3.2. Problems existing in database construction

Summarizing the shortcomings of self-built database construction: (1) The survey found that most databases have stopped updating. Database that is still being maintained also has a large amount of information outdated or resource link failure phenomenon. (2) All the existing self-built libraries still rely on manual resource collection. The new information sources are difficult to be found in time. (3) The subject resources cannot be consistent with the change of subject research direction. (4) Most of visit are restricted by IP address, so the resources cannot be fully utilized. (5) Limited by the technical constraints at that time, the self-built database can only provide simple functions. Network security of system platforms cannot be guaranteed. Many services cannot be continued, such as RSS push function and social services. Based on the old technology, the data of different data types or topics cannot be recorded by unified standard, resulting in that data cannot be transferred and shared. (6) With the continuous addition of new technology, users can obtain resources in a variety of ways, and the old resources cannot be configured to the library unified retrieval platform, which leads to the continuous decline of the utilization rate of self-built resources. (7) Few libraries conduct overall analysis and research on their own resource system, which leads to the lack of overall view on the construction direction of self-built database, and finally makes it become empty talk as an effective supplement to library resource system.

4. TRANSFORMATION STRATEGY DRIVEN BY BIG DATA

The shortage of library funds, the explosive growth of network resources and the free opening of more and more academic resources make the construction of self-built resources promising, while the wide application of artificial intelligence technology provides the feasibility [5]. Combining the experience and lessons of self-built resources construction in the early stage, the formulation of transformation strategies driven by big data can make

full use of strengths and avoid weaknesses in future construction, making it become a real powerful supplement of library resources, which can relieve the shortage of library funds, and guide the service innovation of library.

4.1. Artificial intelligence embedded in libraries

Modern libraries are the center of knowledge management and services. Professor Nelson from the Artificial Intelligence Research Center of Stanford University believes that artificial intelligence is a science about how to represent knowledge and how to obtain and use knowledge [6]. Based on the level of knowledge management, artificial intelligence technology will become an essential technology in libraries. Artificial intelligence expresses knowledge through state space, problem reduction, semantic network, data framework, predicate logic, knowledge process and other technologies, and forms a knowledge map integrating various information forms and interaction modes, which can realize the processing and management of multiple types of data on a unified platform and is conducive to data mining and utilization. It breaks the limitation of data type and subject division in resource collection, and facilitates users to acquire knowledge flexibly through intelligent search.

In the self-built resource transformation strategy, the library can use artificial intelligence technology to customize the resource collection strategy, automatically complete the resource collection and indexing, and use machine learning to constantly revise the resource collection strategy, so as to realize the active discovery of knowledge and ensure the comprehensiveness, accuracy and timeliness of library big data collection. This technology can not only improve the data quality of library self-built resources, but also analyze the shortage of library procurement resources, which is more helpful for staff to control the key direction of self-built resources collection. For another, the library can actively obtain the needs of users through the perceptual intelligence realized by face, image, sound and other perceptual devices and artificial intelligence algorithms. In the transformation strategy, the library is no longer passively accepting the user's demand and suggestion, but can actively analyze the user's demand and improve the construction and service innovation of the library resource system on the basis of this.

4.2. Library big data

Under the mode of big data thinking, the most important link of transformation strategy is the determination of library big data scope. Extensive, diversified and related library data should include four parts: collection resource data, including library books, periodical data, digital resource data, self-built resource data, etc.; library transactional data, including user behavior data, statistical analysis data, business data, etc.; school-related data,

including educational administration data, scientific research data, academic activity information, scientific research project information, community activity information, and curriculum information; network-related data, such as open access resources, academic conferences, major scientific devices, society organization, etc.

4.3. Library smart service

In big data environment, users have higher requirements for library services, so the transformation strategy should keep up with the needs of users. According to the current experience of artificial intelligence in various fields, the transformation of library service should be considered from the following aspects:

From passive to active. By analyzing data related to user behavior, library entry information, borrowing information, and electronic resource usage information, the library uses perceptive intelligence technology to learn users' personalized needs, and uses artificial intelligence technology to organize users' required learning and research resources, so as to provide users with intelligent and active push services.

From complex to simple. The natural language expression, expert system, intelligent retrieval and other technologies of artificial intelligence will change the past resource service mode of library. Users no longer need to write complex retrieval, but put forward their own requirements in the most convenient expression mode, so that a group of related knowledge, visual analysis results, or an information set needed for research and decision can be obtained.

From one side to interaction. Based on artificial intelligence knowledge map, multimedia identification and other technologies, library services will become a process of interaction between libraries and users. In the process of making the request, the platform can analyze the user's preferences through known data and give the most appropriate results according to the user's needs.

The unified data description standard and the unified platform for data management also provide the possibility for the expansion of data applications and the diversified development of service channels. Libraries can provide services to users through Weibo, WeChat, QQ and mobile APP.

4.4. Positioning of self-built resources

Big data technology can provide development optimization strategy for library self-built resources construction. Through the collection, preservation and analysis of the library big data, the visual technology can be used to reveal the guarantee and utilization of the existing resources of the library more intuitively, and the change of the needs of the library users can be predicted more accurately.

By analyzing and mining the hidden value of data association, users' demands can be analyzed, and combining with the actual utilization benefit of database, libraries can allocate resources more reasonably with funds. In turn, the comprehensive allocation of resources and user needs can accurately locate the construction and development of self-built resources, thus saving the manpower and material resources of the library in the process of self-construction. It can also improve the efficiency of resource construction and provide more effective resource service for library users.

4.5. Data collaborative construction

The open sharing of data is the trend of the development of the times, and public access to government information has been in the forefront. Public libraries have also done a lot of frontier work in this field, such as "Genealogical Knowledge Service Platform" launched by Shanghai Library [7]. As the center of knowledge inheritance, management and service, the university library should formulate the strategy of data co-construction as soon as possible, set up provincial or national university big data center with the university library alliance as the lead unit, and plan the self-built data platform to realize the open sharing of data.

From the long-term perspective of data cooperation construction, each school should consider the future integration and sharing of data in the construction of self-built resources. It is necessary to establish a reasonable incentive mechanism for co-construction and sharing, so as to promote the co-construction and development of data in colleges and universities.

The establishment of big data center in colleges and universities will promote cooperation and competition among colleges and universities, stimulate their respective advantages and potential, realize distributed cooperation and develop more diverse characteristic applications.

5. CONCLUSION

Based on the investigation and analysis of self-built resources of university library in Hubei province, this work pointed out that the old university self-built database technology is backward, the maintenance is stagnant, the data quality is intermingled, the platform function is low, the system network security has serious hidden trouble, and the number of databases reduces sharply. As an important supplement to the library digital resource system, the current self-built resources cannot meet their own development needs, and the service mode cannot meet the needs of users. Therefore, the transformation of self-built resources under the background of big data is imperative. This work discussed the transformation strategy of library self-built resources from five aspects: artificial intelligence technology, library big data, library intelligent service, location of self-built resources and data

cooperative construction. The development of library self-built resources can play its important role and provide the best resource support for the healthy development of library resource system and user service only by adapting to the progress of science and technology, innovating and putting into practice.

ACKNOWLEDGMENT

This work was supported by Scientific Research Fund Project of Hubei University Drawing Commission in 2019 (No. 2019 YB11).

REFERENCES

- [1] Chen Si, Xu Junhua and Zhou Yonghong. Investigation and Analysis of the Self-built Database Access Entrance of "985" University Libraries - Based on the Experience of Famous University Libraries in Europe and America. *Research on Library Science*, No.08, pp.46-55, 2016. DOI:<https://doi.org/10.15941/j.cnki.issn1001-0424.2016.08.009>.
- [2] Chen Shanmin. The Research Trend of Characteristic Resources in the Field of Books and Information in China - Based on Literature Measurement and Multivariate Statistics. *Information Science*, vol.35, No.08, pp. 170-176, 2017. DOI: <https://doi.org/10.13833/j.cnki.is.2017.08.062>.
- [3] Han Yuzhe. Characteristics Analysis and Enlightenment of Foreign Digital Humanities Projects. *Library*, No. 11, pp. 18-23, 2019. DOI: <https://doi.org/10.3969/j.issn.1002-1558.2019.11.004>.
- [4] Access, Wuhan University of Technology Material Composites New Technology Information Portal, 10 July 2010, <http://atmsp.whut.edu.cn/>
- [5] Wang Shiwei. Three Stages of the Development of "Whole-process Intelligence" in Public Libraries During the 14th Five-year Plan. *Library Construction, Network Starting Time: 2020-09-09-09 10: 37: 39*.
- [6] Ma Shaoping and Zhu Xiaoyan. *Artificial intelligence*. Beijing: Tsinghua University Press, p.2, 2004.