

The New Development of Digital Libraries in China

Wei Liu

wliu@libnet.sh.cn

Shanghai Library

Shanghai 200031, China

Abstract

National level digital library infrastructure has a high priority to Chinese policy makers, just as the same important as the internet communication system to the national information infrastructure. According to the literature and the statistics from periodical database, the number of papers on digital library presents exponential growth in recent years. It reflects that the integration and services of digital collections have become the main issues in most of Chinese large libraries, including public, academic and special libraries. This paper reviews and highlights all the main efforts on the research and construction of digital libraries in China, introduces the evolution of the developing policy of digital resources, the achievements about the digital services, and the trends of digital library development in China, as well as the challenges on technology and policy aspects at the national level.

Keyword: digital libraries, china, metadata application, projects

Introduction

Like in the United States and Europe, while the research and construction of digital libraries are prioritized as most important component in the national information infrastructure, the term “digital library” in China has no unified understanding or definition. It varies with different situations and circumstances. In general, digital libraries have presented themselves as two major aspects in Chinese publications, one as a research area involved with computer specialists and content service providers who focus on technology improvement on digitizing, organizing, distributing, preserving, retrieving and searching, while the other

as practical application projects aim to build up digital collections or provide digital library services which are involved with librarians and content owners.

China is a large country with a great imbalance in the distribution of resource and population. Its GDP (Gross Domestic Product) per capita is less than 1000 USDⁱ in 2002. But the information infrastructure of China has achieved greatly in recent years. According to the biannual survey by CNNICⁱⁱ, there are 371600 websites with 157 million web pages, 83 thousand online databases and 2878GB information within the cn domain by the end of 2002. There are more than half of the 59 million internet users lives in the capital and a dozen coastal cities, which is corresponding with the distribution of universities and scientific research institutions, as well as the industries which are located around the coastal regions. This reflects varying levels of librarianship and information literacy in China. The research and development of digital libraries is usually based in Beijing - the capital city - and other big cities that are located in the coastal areas.

The majority of funding for digital library projects comes from the central government, mostly from the national foundations for science and technology R&D (under the name of “863” and “973” High-tech Development Program) and the funds earmarked for construction of national information infrastructure. National Social Science Foundation is also a national funding source. A small part is from the provincial government, institutions and affiliated libraries. One new development is that digital library research institutes have been established by some big universities, large public libraries and corporationsⁱⁱⁱ to deal with the increasingly common requirements for the

integration of digital resources and the digitization of traditional materials.

Brief History

Chinese libraries have been paying close attention to the application of new information technology and the development of digital libraries in western countries since the beginning of last nineties. Many Chinese librarians heard the word “digital library” for the first time in a special session: “Digital Libraries, Technologies and Organizational Impacts” during the 62nd IFLA General Conference in 1996 held in Beijing. But the first project named after “digital library” in China is initiated in 1994, which was consisted of three sub-projects and cooperated between IBM and Qinghua University, Fudan University and Chinese National Petroleum Corporation (CNPC) respectively. The project aims to build up stand-alone mass data processing systems for images and multimedia information resources.

The “National Pilot Digital Library Project (NPDL)”^{iv}, which was initiated by the Ministry of Culture to State Planning Commission in 1997, was the first national level digital library project in China. The research partners include National Library of China and five other large public libraries. Before the project, there were few papers on digital library giving only brief introduction to projects from abroad, for example, the projects of Digital Library Initiative Phase one from the United States.

Since 1998, apart from the national library and public libraries led by the Ministry of Culture, there have been many research and development projects related to digital library launched by many universities under the Ministry of Education

(formerly the State Education Commission of China), and the science and technology research institutes, which are led by the Ministry of Science and Technology (formerly the State Science and Technology Commission of China), and the Chinese Academy of Science. Chinese Vice Premier Li Lanqing commented that “Digital library will be the trend of libraries in the future” as he visited National Library of China in October 1998. After that, many key programmes or projects were initiated, including the most significant and costly project by now (about 1.2 billion RMB): the National Digital Library Project of China. And the digital library construction has become a hot topic in China.

Literature analysis

A literature search was conducted on published papers collected from CNKI and Chongqing VIP Database vi. The China National Knowledge Infrastructure (CNKI) is a digital library which comprises seven full text databases: Journals, newspapers, monographs, dissertations and conference proceeding, and two specializing in the medical and education fields. The Chongqing VIP Information Services Company produces a comprehensive periodicals database.

According to the statistics from CNKI, there are more than 1600 papers published with the keyword “digital library” from 1994 to the first quarter of 2003 in Chinese periodicals. The search returned hits with the highest count of 689 in 2002 which presents an approximatively exponential growth from the beginning of 1994, as shown in table 1. The hits returned from VIP database which includes more periodicals has a bigger number but seems to tell the same story (see table 2).

	1994-1997	1998	1999	2000	2001	2002	2003	1994-2003
Ti=“digital library”	19	19	54	127	318	475	127	1139
Kw=“digital library”	20	23	62	175	473	689	162	1604

Table 1: the growth of digital library related papers from 1994 to 2003 (from CNKI)

	1995-1998	1999	2000	2001	2002	2003	1995-2003
--	-----------	------	------	------	------	------	-----------

Ti="digital library"	57	54	197	423	661	331	1723
Kw="digital library"	80	80	275	646	975	499	2558

Table 2: the growth of digital library related papers from 1994 to 2003 (from VIP database)

Comparing with the several thousand papers issued every year by Chinese library professional

periodicals, the increasing in the number of papers related to digital library is extraordinary. Table 3 gives a brief content analysis (by sub-category) of the digital library related papers.

Subject	CNKI entries	Percentage
Digital library construction	433	27%
Digitalization or digital resources development	154	10%
Intellectual Property	111	7%
Technology R&D	109	7%
Information services	51	3%
Digital resource Sharing	48	3%
Business model or service model	40	2%
Library automation	38	2%
Digital library architecture	28	2%
User study	27	2%

Table 3: content analysis of digital library related papers

Although the majority of the papers are review and discussion articles, the entries in each category reflect the research interests within Chinese digital library circle.

Digital library projects

From the mid-nineties of last century, quite a number of projects named with "digital library", initiated by national or provincial institution. Some of them are research projects while majority of them are resource digitalization or digital library construction programs. Listed below are most significant projects and programs in recent years.

Digital Library Research Projects

Digital library projects in National 863 Program

The National High Technology Research and Development Program (863 Program) was launched in March 1986 with the aim of enhancing China's international competitiveness and improving China's overall capability of R&D in high technology. Digital library is one of the subject

topics within information technology sector, which are managed by the Ministry of Science and Technology. There are four digital library related projects initiated in 863 Program after 1998.

1. Strategic Research on the Development of Chinese Digital Libraries
2. Research on characteristics-based multimedia information retrieval system^{vii}
3. Knowledge Network: Digital Library System Project^{viii}
4. Chinese Digital Library Application System on China Advanced INfo. Optic NET (CAINONET)^{ix}

Chinese Pilot Digital Library Project (CPDLP)^x

In May 1996, the National Library of China, together with the major public libraries in the country, including the Shanghai Library, Liaoning Library, Nanjing Library, Zhongshan Library and Shenzheng Library, submitted a proposal for China's Experimental Digital Library Project to the State Planning and Development Committee. This project was instituted in 1997, and designated one of the state key science and technology projects. It was completed in May, 2001.

CPDLP aims to establish an experimental digital

library system with the collaboration, resource sharing and unity of all participants. It operates in a distributed environment formed with standardized and multi-type resources. It is the first digital library project in China and the first to give a solution to issues in developing digital libraries in the country. The project needs further improvement to be put into massive use in developing digital libraries in China. The solution is made up of three aspects: the digital resources construction system, the resources store and management system and the services system.

The Establishment of Standards and Specifications for Digital Library Construction in China (ESSDLC) project

In September 2002, the project “Establishment of Standard and Specifications for Digital Library Construction in China” was proposed to the Ministry of Science and Technology by the National Science and Technology Library, National Library, China Academic Library and Information System (CALIS), Library of Beijing University, Library of Chinese Academy of Science, Institute of Scientific and Technological Information of China, and Shanghai Library etc. This is a real joint project in which the members are come from all the leading organizations related to digital recourses. Most of the each partner has led one sub-project and there were six sub-projects started at the end of 2002., three of them started at the end of 2003.

The goal of the ESSDLC is to generate a series of national standards and specifications regarding to the construction of digital libraries in China. It is also a feasibility study to realize in what degree the necessity is, for each of these standards and specifications in the sub-projects. The ongoing project has already issued 35 reports in its website^{xi} and some of them were published in professional journals.

Chinese Metadata Standard project

In most of the digital library related projects which

sponsored or initiated by National Library of China, the research and compiling of metadata standard and specification has taken a very important role. “The First Conference on Construction and Sharing of Chinese Documents” which held in Beijing on the June 2000 pleaded for the “Chinese Metadata Standard” to fulfill the needs on Chinese information resources construction. The conference urged NLC to lead the work on drafting the Chinese Metadata Standard. The Ministry of Culture approved the project proposed by NLC. On June 2001, based on the tracking of the foreign research and its own experiences for a long period in this area, and integrated with the characteristics of Chinese document, NLC has finished the draft version of “Chinese Metadata Standard”^{xii}.

The Standard adopted an overall framework which followed the reference model presented by OAIS. As for metadata element set, it refers to the similar projects from the Library of Congress, the National Library of Australia, and the Cedars project, the NEDLIB project and Dublin Core Metadata Initiative.. The core metadata set contains the majority of Dublin Core Metadata Set, but adds a lot of supplementary elements for the description of intellectual properties and abstract format of digital resources. The draft version of March 2002 consists of 25 core elements, their qualifiers and the encoding schema with DTD and RDFs. It has been considered to have the suitable coverage on characteristics of Chinese Documents with adequate descript ability, functionality, operability and simplicity.

Digital library construction program

The China Digital Library Program (CDLP)

As a major successive achievement of the CPDLP, in July 1998, the National Library of China brought forward its Chinese Digital Library Program (CDLP) proposal and it was approved by China Government. Much like National Digital Library Project in the United States, the main goal of the CDLP is focus on composing a digital library of reproductions of

primary source materials to ensure better preservation and study of Chinese culture heritage. Begun in 2000, the project proposed a five years plan to digitize a huge amount of its collections, including books, images and video audio collections. With technical achievements by CPDLP, CDLP set up a national-wide platform to enhance the access to digital repositories yielded by this project. Lead directly by the Ministry of Culture, CDLP takes the advantages of been a national level project. A consortium which has 117 members was established by the end of 2002, under a practical rule of “unique planning, standards ruled, team work and resource sharing”.

According to these objectives, the major tasks of the program involve the creation of digital repositories, the construction of digital library software/hardware infrastructure, the development of application systems, the shaping of standards and specifications structure, and training of professionals.

National Library of China Phase II & National Digital Library of China Project

Approved by the State Development and Reform Commission of China, National Library of China plans to implement the National Library of China Phase II and National Digital Library of China Project. Actually this project is a part of successive achievement of CDLP, including a bibliotheca containing 12 million books with nearly 78 thousand square meters, reading rooms with 2000 seats, a National Digital Library and its accessory facilities, at a total budget of RMB 1.2 billion.

National Culture Resources Sharing Project (NCRSP)

Also lead by Ministry of Culture in China, National Culture Resources Sharing Project (NCRSP) has lots of ties with CDLP. The principal objective of NCRSP is to collect and digitalize various type of large amount culture content through out the country, make them available for domestic masses by nation-wide network. Unlike the CDLP, which

focus on technology, standards and application development, NCRSP pays fully attention to content creation, network construction and services accessibility based on the technologies and resources that CDLP carried out.

This project is a major effort conducted by government to improve the unfavorable information resource sharing situation in the country. By the end of 2005, NCRSP intends to establish a network with three tiers which have a national level center, more than 30 provincial information center and more than 5000 information centers scattered in counties, towns and communities. With the network, NCRSP can ensure the connection and availability with the masses. A union catalog with the culture content in libraries, museums, art galleries and art institutes over the country being plan to made at first, this will take advantages with the union content creation, handle system and resource discovery. By the end of 2005, NCRSP plan to finish a digitalization of one million books, one thousand local operas, one thousand music production, one thousand art works and one thousand antique items. The digital copies of Chinese culture heritage will be accessible to as much as people in China as soon as the project accomplished.

National Technological Library and Information Centre^{xiii}

Led by the state Scientific and Technological Ministry and approved by the State Council, the National Technological Library and Information Centre, was established in June 2000. Aims at meeting the needs of technological information by building a visual centre of technological information including science, engineering, agriculture and medicine, the center is actually a digital library of scientific and technological resources. The participating institutes are the China Academy of Science Library, the China Scientific and Technological Information Research Institute, the Mechanical Industry Information Research Institute, the Metallurgical Industry Information and Standard Research Institute, the China Chemical

Industry Information Center, the China Academy of Agriculture Library and the China Academy of Medicine Library. There are totally 7 libraries and Institutes.

The Goal of Development is to create a national scientific and technological resource collection and service consortium. The general principle on which the consortium is based is that each institute will process data and provide services separately, but will cooperate in establishing a central database for the uniform information retrieval. The center will provide users with secondary document retrieval services and primary document delivery services through Internet. Any Internet users can have free access to the secondary document retrieval service provided by the center. Registered users can also request the center to deliver primary documents in various ways, such as by email, fax and the postal service.

Chinese National Science Digital Library (CSDL)

As a part of the Knowledge Innovation Project of Chinese Academia, CSDL has been started at Dec. 2001. Primary goal of CSDL is to provide better support to scientific research and knowledge service with digital library framework and technology. Main tasks of CSDL is to build integrity information resources and management system all over the Chinese Academia, promote and speed value-added information exchange and knowledge delivery service under the network circumstance, and make the digital library to be an extension to every scientist's desktop in China.

CSDL not only purchases lots of important commercial electronic resources, but builds up its own Chinese scientific literature databases with a amount of more than one million records, which covers major topics like chemistry, biology, physics, mechanics, optics, mathematics, astronomy, geography, computer science as well. Further more, the access to the content in the CSDL does not limited to just searching its bibliographic database and electronic resources. With the capabilities of newly built inter-library loan and

document delivery service system, CSDL cover its service across the whole Chinese Academia across the country. CSDL can also provide more than ten-thousand kinds of full text of foreign scientific magazine and other collections, disciplinary portal of special scientific domain and virtual reference services.

China National Knowledge Infrastructure (CNKI) Base^{xiv}

Launched in June 2000, the China National Knowledge Infrastructure consists of over 70 special databases, covering theoretical research, policy research, laws and regulations, work discussions, current news and affairs and other related information.

The CNKI is based on the China Periodical Network Full Text Database and the China Major Newspaper Full Text Database created by the Qinghua University. But there will be some Chinese and foreign databases such as "China Laws and Regulations Database", "China Trade Standards Database", "Chinese Yearbooks Database" etc integrated to the services. The CNKI knowledge base provides online subscription services with a central web site and 80 mirror sites countrywide.

China Academic Library & Information System (CALIS) Project

Led by the state Educational Ministry, the China Academic Library & Information System (CALIS) was approved by the state Project and Development Committee and launched on November 1998.

The CALIS project was a library-networking project at the beginning, aiming at web information resource sharing via CERNET (China Education and Research NETwork) for faculties of colleges and universities across the country. With the web-based technology of distribution, academic libraries should make full use of their superiority in digital resources and speed up the development of their databases. The CALIS project makes use of much in methods that foreign libraries take,

integrates new achievements of technology progress and features for its online cooperative cataloging, the databases originated locally, the customized OPAC system and the interlibrary loan system.

Discussion and Conclusion

With the booming of knowledge economy, information resources are being mined and widely used. As the “thesauruses of knowledge”, libraries will follow the trend of social and economic development to realize the digitization of literature resources and networking of service means, with the push and support from modern information technology. But in China, the concept of “digital library” arises with the growing of the internet bubbles. Researchers and practitioners of digital library have promised too much to decision makers and stakeholders, so that they can get the financial support or approval for their projects from the central or provincial government. When it is still a R&D area in the United States, the large scale national level implementation project is about to embark in China. It’s a little bit dangerous to build digital library buildings in such a fast changing basis of Information technology.

For most of the libraries in China, the first challenge and motivation of digital library construction is from the electronic resources flooding into libraries without control. One third of budget used for digital collection development in a few research and academic libraries in China. Traditionally, catalogue is the image of library collections and the main entrance for readers to access. But more and more “virtual” digital collections provided by publishers or information brokers can hardly integrated into library OPAC system. Libraries have to deal with three kinds of digital resources: digitized traditional collections, leased databases or born-digital materials, and free-access materials. Library collections have no longer bounded within library walls. There is no sole portal for a library providing to its readers in the digital age.

Chinese government has made a great effort to

promote digital library construction. But there are still some big issues to be dealt with:

Intellectual Property: China has established a modern intellectual property system in just two decades; as a result, awareness of intellectual property rights remains underdeveloped in society at large. In most cases, libraries are not the owner of their collections. So far materials in public domain are always the first choice to digital library projects. But the term “public domain” remains unclear in practice. There always have been a large number of things to do before the start of digitization and providing “fair use” services to digital library users. Digital copyright will always be a heavy burden for digital library practitioners till the reasonable intellectual property framework for digital assets perfectly settled down. The intellectual property framework is challenged by the coming digital society. Digital library technology has to provide a flexible and extensible IP management framework in applications.

Cost for digitization: With a history of 5000 years, ancient collections in Chinese libraries are the embodiment of the wisdom of the past. This cultural heritage is an important part of human legacy and can contribute immensely to the digital mainstream of present world. This is also considered to be the best way to preserve the culture heritage. But libraries themselves are not publishing houses. The high cost of digitization is not affordable just by libraries. Chinese government has made a great effort to establish and support the national information infrastructure. The multi investment mechanism should take into account within the national policy for digital library development.

Cutting-edge Information Technology: There is such an opinion in China that no necessary to worry about technology. You can get any technology for digital library when you get money. But that is not true. Digital library consists of all kinds of cutting-edge information technology without special favors. It seems we should be the master of technologies but technologies lead to our modern life, no exception for the future of digital library too. China has a lot of technology achievements in

multimedia retrieval, full text procession, distributed computing, etc., but as a whole, China is lack of systematic research and development on digital library technologies.

Standards and Specifications: As a digital resource and service providing system on the internet, the digital library has to establish and follow a set of standards and specifications about digitization processing, resource description and organization, system interoperating and services. It should adopt all kinds of existing related standards and specifications of content encoding, data communication, intellectual property protection, system security, management, service and operation, so as to assure the usability, interoperability and sustainability of the resources and services of digital library. But there will be a long way to go to conclude a necessary technology standards and specifications system to smooth the construction of digital library in China.

References

- [1]. Xia Cuijuan "Reviews of the Researches on Digital Library of China 2001-2002"
<http://www.library.sh.cn/libnet/sztsg/fulltext/reports/2003/DLOverview2002.htm> (viewed 1/4/2003)
- [2]. Wang Qiyun "The Investigation and Analysis on Thesis of Digital Library in Recent 6 years" The

Modern Technology of Library and Information Science 2002.2

- [3]. Xia Xu, Zhang Chunhui etc. "The Presence and Development of Chinese Metadata Research" Proceedings of International Conference on Chinese Metadata Application. 2001. Shanghai
- [4]. Arms, W.Y."Digital Libraries", MIT Press 2000
- [5]. Liu Wei, etc. "Introduction to Digital Library". Shanghai Science and Technology Document Press 2000. 12
- [6]. Zhu Qiang "Pre-Requirements of the Digital Library" Journal of Academic Libraries (Beijing) 2000 no.1
- [7]. Sun Chengjian Liu Gang "The Outset and Development of the Digital Libraries in China" the White paper of Digital Library/Computer World 1999. 6
- [8]. Wang Bing "Digital library: definitions, Impacts and Related problems" Bulletin of China Society of Library Science (The Journal of the Library Science in China) 1998 no.6
- [9]. Zhang Xiaolin etc. the Trends on Digital Library Standardization, the Journal of Chinese Librarianship 2002.6

ⁱ According to the World Bank Report, see <http://www.worldbank.org.cn/>

ⁱⁱ See <http://www.cnnic.net.cn/tj/rep2002.shtml>

ⁱⁱⁱ So far there have been several digital library research institutes established by Beijing University, Tsinghua University, Shanghai Jiaotong University, Shanghai Library and Shanghai Changjiang Computer Group.

^{iv} See: <http://www.library.sh.cn/libnet/sztsg/fulltext/reports/1998/metadata.html>

^v See: <http://www.nlc.gov.cn/dloff/>

^{vi} Based on the data collected from CNKI (<http://www.cnki.net>) and Chongqing VIP Database (<http://202.119.47.6/>), viewed in 2003.12

^{vii} See http://www.nlc.gov.cn/dloff/scientific6/sci_2.htm (2003/12/25)

^{viii} See

http://www.nlc.gov.cn/dloff/scientific6/sci_5.htm (2003/12/25)

^{ix} See

http://www.nlc.gov.cn/dloff/scientific6/sci_6.htm (2003/12/25)

^x [The alias is China's Experimental Digital Library Project](#), see

http://www.nlc.gov.cn/dloff/scientific6/sci_7.htm (2003/12/25)

^{xi} See: <http://cdls.nstl.gov.cn/cdls2/w3c/>

^{xii} Can be downloaded from:

<http://www.cdi.cn/download/dmds.pdf>

^{xiii} <http://www.nstl.gov.cn>

^{xiv} <http://www.cnki.net/>