THE INTEGRATION OF LIBRARY, NETWORK, AND TELECOM: DATA AND ITS SEMANTIC EXPANSION

Dasheng Wang[1] * *and Lei Liu*[2]

*[1] *Library of Beijing Information Science & Technology University, wds84192@biti.edu.cn*

[2] The Chinese National Committee for CODATA, Computer Network Information Center, Chinese Academy of Sciences, 4, South 4th Street, Zhongguanchun, Beijing 100190, China,
Email: liuleirobert369@163.com or Email: liulei@codata.cn

ABSTRACT

The concept of information is being expanded, which extends the scope of libraries' collections; the application of the theory of electromagnetic-field, and the combination of telecom, computer with consciousness of the mankind, have realized information spread online; The essence of retrieval and reading in library decides libraries' integration with telecom and network. The integration's appearance, essence, and means are discussed. Data is more and more important during the course of the integration and information service. The content of metadata goes beyond the pure scope of numerical value and number range. The semantic expansion is happening in the concept of data.

Keywords: integration; library; network; telecom; data; metadata; database

1 INTRODCUTION

The connection of telecom, computer platform and human consciousness has realized the network transmission of sound, words and pictures. The essence of libraries satisfying people's retrieval and reading requirement decides the inevitability of its integration with telecom and network. The integration enables people—anywhere where there is telecom and network—to do what can be done only in libraries formerly. We can see the course of the integration from facts and data. Data is playing a more and more important role during the course of digitization of information resources and information service. The use of metadata and database management system in libraries has caused semantic expansion of the word of "data".

2 LIBRARY AND INFORMATION

Library in ancient and modern times was only related with paper-based information. Books, periodicals, newspapers and other written materials are carriers that convey something for people to hear and see. Modern library broke the limit of paper-based information with the appearance of sound information and paperless reading. According to the theory of electromagnetism-field（Wang Xianchong,1986, "Information is the combination of signal (or state) from universe and medium substance. Information is from substance, but is not the substance where it is from. It is a new substance that is the combination of medium substance and source substance. Such a change happens in the twinkling of an eye (Wang Dasheng & Liu Bin, 2005)". For this reason, where there is substance movement there must be information. As long as the universe exists, information interaction won't stop. Ever since human society came into being, information produced by interactions between people and nature, between people and society, between people and people, was called social information. Mankind's becoming aware of and use of information in objective world begins with instincts of seeing, hearing and touching. In primitive society ape-man's all behaviors like hunting, picking, etc. took place on the basis of receiving some kind of target information. With the development of society, people created in their working process linguistic information used for exchanges, communication and cooperation. Human language replacing animal cry is a revolution in the history of information development. Information formed by printing and processing with paper as its carrier is called literature information, which is the component part of social information. The appearance of words made people have deeper understanding of

information, and then newspapers, magazines and books were invented. According to provisions of UNESCO, bound presswork with more than 49 pages are called books, while 5-48 pages booklets, less than 4 pages scattered materials. Books and reference materials are information that were chosen by people, were thought useful, could be kept for a long time, and were called knowledge. Library is an establishment and building where people collect, put in order, store literature information; information involved in a library is the more essential part of all social information. The human cultural heritage and scientific knowledge as literature information that are classified and stored in libraries for people to search, consult and utilize. China has published hard copy literature including 6,600 million books with 220,000 kinds, 43,900 million newspapers, and 2,900 million periodicals (NBSC, 2008). Library collections also include sounds, videos, etc, within which sounds is an important part.

3 LIBRARY AND NETWORK

Seen from history, books were produced more than 3000 years ago, newspapers and periodicals more than 400 years ago, broadcast more than 200 years ago, TV more than 100 years ago, and internet more than 10 years ago. Information technology is the high-tech that develops fastest, is the most pervasive, and is most widely used. The major feature of information society is to use network as its communication means. According to statistic data published by the forum on Internet management of UNESCO on Nov.12, 2007, the total of Global netizens had markedly increased in the last 10 years, from 70 million to 1,200 million during 1997-2007. In Jan.2008, both U.S. and China have more than 210 million netizens, while China follows as the second with only 5 million lesser than U.S. The third is Japan, having 86 million netizens, and from the fourth to the tenth are respectively Germany, India, Brazil, Britain, Korea, France and Italy. The condition of network development in China in last 10 years, see Table 1:

Table 1 10 years' development of Internet in China（1997.10.31-2007.1.23）

Issue time	Total netizens (ten thousand)	Computers online (ten thousand)	CN domain Names (ten thousand)	Broadband Users (ten thousand)	Dial-up Users (ten thousand)
2007.01.23	13700	5940	180.3393	9070	3900
2006.07.19	12300	5450	1 19.0617	7700	4750
2006.01.17	11100	4950	109.6924	6430	5 1 00
2005.07.21	10300	4560	62.2534	5300	4950
2005.01.19	9400	4160	43.2077	4280	5240
2004.07.20	8700	3630	38	3110	5 1 55
2004.01.15	7950	3089	34	1740	4916
2003.07.21	6800	2572	25	980	4501
2003.01.16	5910	2083	17. 9	660	4080

2002.07.22	4580	1613	12. 6	200	2682
2002.01.15	3370	1254	12. 7	No statistics	2133
2001.07.17	2650	1002	12. 8	No statistics	1793
2001.01.17	2250	892	12. 2	No statistics	1543
2000.07.27	1690	650	9. 9	No statistics	1176
2000.01.18	890	350	4. 8	No statistics	666
1999.12.05	400	146	2. 9	No statistics	256
1998.06.30	1 17. 5	54. 2	0.9415	No statistics	46
1997.10.31	63	29. 9	0.4066	No statistics	25

Till 30 June, 2007, China had 160 million netizens, and popularization rate reached 12.3%. See details in Figure 1 below:

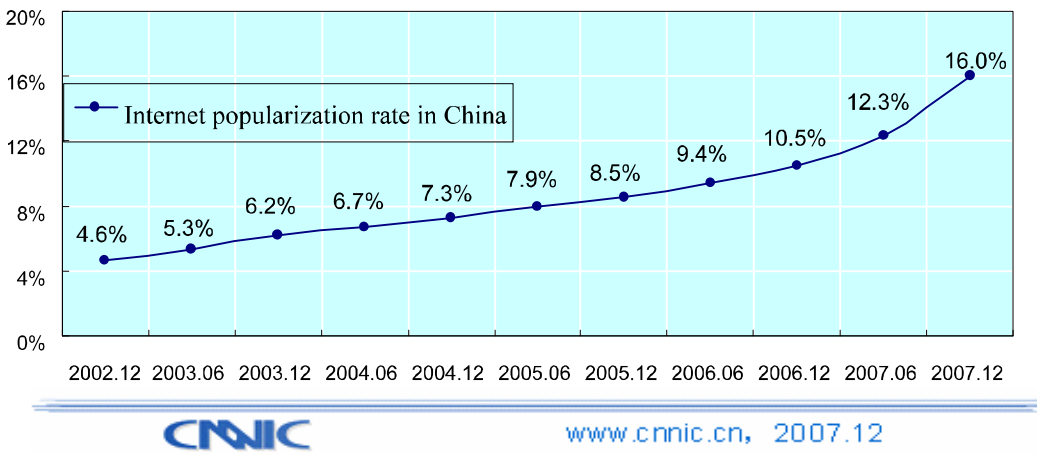

Figure 1. Internet popularization rate in China in 6 years (2002-2007）

Till 31 Dec. 2007, China had 210 million netizens, among whom broadband users were 77 million. Averagely the internet popularization rate is 16% in Chinese mainland, 19.1% lower than the average level in the world (CNNIC, 2007). Chinese netizens get online for averagely 16.2 hours each week; 62.1% of the netizens choose to get online at home, 43.3% in work unit, 21.8% in school, 17.3% in cybercafé or coffeehouse,0.7% in public libraries, and 0.9% using mobile internet providers. Home is the uppermost place where netizens get online. Readers have been shaping the habit of reading online, individual electronic reading terminals had exceeded 150 million, and 1/4 of the readers are reading digital publications, while the number of such readers is increasing by 30% each year. Netizens in France

get on-line for 13 hours averagely each week, while netizens in Britain and Spain 11 hours, netizens in Italy 8 hours, and netizens in all Europe 10.25 hours each week. Seen from facilities, there are more than 6,000 million people in the world, more than 15,000 million microcomputers. Facilities are becoming more and more intelligentized. According to the forecast of IDC (International Data Corporation), in 2001-2008 PC market in China will keep a complex growth rate of 15.5%, and will surpass America and Japan to become the greatest market in the world. Network mainly affects library in the following aspects:

3.1 Digital library

Network has brought up new generation of "digital library". In 1994 the US Government invested 22 million U.S. dollars to carry out its Digital Library Initiative in 6 universities including Carnegie Mellon University and so on. Ever since in the world there have been hundreds of libraries that are using information technology to transform hard copy information into digital information, then into online information. On June 7, 2007, twelve famous universities in America joined Google's book scanning plan, and they contributed one million books for Google to digitize, aiming to build an open-access digital books warehouse for teachers, students and the public to use in any time. In Dec. 2007, the Chinese contractor of the Universal digital Library (http://www.ulib.org/) declared in Hangzhou, after 7 years' construction, this Universal Digital Library jointly constructed by China, America, India and Egypt had possessed more than 1.5 million books. As the largest commonweal digital library in the world, it facilitates university libraries to provide free service for all users using any language anywhere anytime. Up to now the number of e-books in more than 20 languages in this library has accounted for 1% of the total of all books in the world, and most of the digital books are in Chinese and English. During the Universal Digital Library construction, the Chinese part (CADAL item, www.cadal.zju.edu.cn) is the fastest. Till now China has digitized 1.07 million books, accounting for 70% of all books in the Universal Digital Library. The website of CADAL has attracted users from more than 80 countries, and its daily click times have reached 300,000; 4000 e-books are being glanced over each day (Zhang Le, 2007). In last 10 years (1997-2007), five kinds of national digital library projects had been finished: National Digital Library (NDL), China Academic Digital Library and China Academic Digital Library & Information System (CADL and CALIS), China Science Digital Library (CSDL), National Science and Technology Library (NSTL), Party School and Army Educational Institutions Libraries (PSAEIL). These libraries are backbone of Chinese digital library (Zhan Furui, 2008).

3.2 Online reading

China Academic Library & Information System (CALIS), whose online reading is widely used, has built 4 national literature information centers on arts & science, engineering, agriculture, medicine in Beijing, which forms the first layer of resource guarantee system of CALIS, mainly playing the role of guarantee base of literature information. Within CALIS the national literature information centers on arts & science and engineering are separately set up in Peking University and Tsinghua University. CALIS' online reading function can basically satisfy requirement of readers in China. By 2006, Chinese Journal Full-text Database (CJFD) embodied 6990 journals; National Knowledge Infrastructure (CNKI) embodied 7582 journals (except those of science popularization and annals). Vip Information (www.cqvip.com) has embodied 400 Chinese newspapers, more than 8200 Chinese journals, and 5000 foreign journals; By Oct. 2006, Vip Information has embodied 16 million articles, with an increase of 2.4 million articles each year. Vip Information has 2000 fixed users, and its website click rate has reached 200,000 times daily. Wan Fang Data (iLib) has embodied 5607 journals. One of its characteristics is having also neatened some information about patents, standards and business. Dragonsource Group com Inc. (www.qikan.com.cn) has employed 1600 kinds of humanities journals covering current political situation, party affairs, management, finance and economics, literature, art, philosophy, history, society, popular science, military, education, family, physical education, leisure, health, fashion and occupation, etc.

3.3 Internet Publication

Internet publication means to edit, make, issue, spread and sell e-books for publishing houses or other circulating media on Internet, which include independent network edition and that corresponds to printing edition. By 2007, Google had acquired 1 million books from publishing houses and provided searching service for its users on Internet, and more than 10,000 publishing houses from the world have joined the Google project of searching for books (Chen Xin, 2008). In recent years, Chinese readers' habit has been changing; the sum of people with the habit of reading online account for only 3.7% of that of all readers in 1998, 18.3% in 2003, and 27.8% in 2005. Online reading rate is dramatically increasing, with an increase of 7.5 times in 7 years from 1999 to 2005, with an average increase of 107% per year. Reader's demand leads to the increase of e-book market. E-book publishing quantity had reached 148,000 kinds in China till April 2005. By Dec. 2007, there had been 300,000 kinds of e-books published in China. Now more than 160 publishing houses are synchronously publishing e-books and hard copy books. Most libraries have been using e-books, and totally more than 1,900 libraries in the country are using e-books. In October 2007, China Digital Newspaper Laboratory publicized the first batch of 45 items in 9 categories, among which newspaper websites, application items on mobile phone newspapers and multimedia digital newspapers & periodicals are the three categories that are submitted most. More than 800 multimedia newspapers have been put online in China. Network publishing has the characteristic of "five no": no printing cost, no physical distribution expenses, no overstock, no distribution expenditure, and no readers' transportation, which makes network publishing have very huge potentiality(Wang Dasheng, 2006). Libraries and publishing houses have the relation of alliance and partnership, the change of publishing modes makes libraries have to transform collection of printed matter into collection of digital resources, and adjust their service modes (Wang Dasheng, 2006). China is pushing on 8 digital publishing projects: 1. R&D of digital multimedia; 2. National digital complex publishing system; 3.Chinese character library construction; 4. National knowledge resources database publishing project; 5. National cartoon promotion project; 6. China ancient books digitizing project; 7. R&D of copyright protection technology; 8. China digital culture communication project (Yan Xiaohong, 2007).

4 LIBRARY AND TELECOM

People had deeper understanding to information when they discover electromagnetic-field, then they invented "telecom", which is the high efficiency and reliable means to spread information. Telecommunication doesn't use air or instinct, but use electromagnetic-field, including wirephoto, radio, light wave and so on, as medium to spread Information like characters, sound and pictures, etc. Specific telcom manners include telegraph, telephone, broadcast, television, optical fiber communications, satellite communications, computers and internet. The use of the theory of electromagnetic-field in the realm of communications is an important revolution of the method of information spread. Experiments indicated that information was spread by electromagnetic-field in the vacuum state, and communications between people on the earth and astronauts by telephone is enough to explain the existence of electromagnetic field and the effect of telcom.

4.1 Mobile phone

Telcom industry has been rapidly developing in recent years. The number of global mobile phone users is increasing by 1 million each day. By Jun. 2007, there are 3,000 million mobile phone users in the world, but this is just the beginning of the popularization of mobile technology. In China, the total of telephone users had reached 912,730,000, among whom the total of fixed telephone users is 365,450,000, and mobile phone users547, 290,000 in December 2007(NBSCC, 2008). The national short message transmission totaled 5,350,800 million pieces from January to November 2007, with an increase of 37.5% than that in the corresponding period last year; In Chinese mainland fixed telephone popularization rate was 28.3% while mobile phone 39.9% in 2007; in Chinese Hong Kong, the

mobile phone popularization rate was 118%, which was first in the world. The 3rd generation communication technology (3G for short) makes transmission of sound and data obviously faster. 3G can deal with transmission in various forms such as pictures, music and video frequency, and provide services like webpage browse, etc., for instance, DAB（Digital Audio Broadcasting）mobile phone. The development of telephones finally results to the birth of "mobile phone library", which has come true in Japan where readers can use mobile phones to read popular writers' works; Mobile phone library includes wireless music (mobile phone rings, mobile phone music, etc.), mobile phone games, mobile phone readings (newspapers, novels, blogs, etc.), mobile phone technical platform (searching engines, two-dimensional bar code, etc.). According to CNNIC(2007), by June 2007, China had had 55,640,000 netizens using wireless terminals, 44,300,000 netizens using mobile phones as their way of access, and others using notebook computers. The integration of libraries and telecom will make mobile phone library develop faster.

4.2 Broadcast

Sound is spread by broadcast, and people receive sound by hearing. Sound is part of library storage. The first broadcasting station in China is established by an American named Ausban in Shanghai in Jan. 1923, while Ha'erbin broadcasting station is the earliest station set up by Chinese in 1926. By the end of 2007 there had been 263 broadcasting stations, 287 television stations, 1993 broadcasting and television stations in Chinese mainland, and 1933 broadcasting channels with a coverage ratio of 95.4 %(NBSC, 2008). There are 189,000 satellite broadcasting receiving stations in Chinese mainland. These stations are working in 40 languages facing the country, while 43 languages facing overseas. On average each station in China broadcasts 20,317 hours everyday (Wang, 2004). China is a country whose number of railway passengers is the biggest in the world annual transported passengers are counted with 1000 million. Each train is equipped with broadcasting system. On Jan.3, 2008, the Ministry of Public Security of China announced that in China there are 159,777,589 motor vehicles; by Dec. 2007, there are 56,967,765 motorcars. 107087137 Motorcar drivers; The number of people who listen to broadcast in cars is rising; The integration makes broadcast as the way to spread books and data to readers, which expands readers' hearing space.

4.3 Television

Watching TV is a kind of reading behavior in an extended sense. IPTV (Internet Protocol Television) can transfer most internet functions into TV. Currently four service providers having acquired operation licenses in China are Shanghai Media Group (SMG), China Central Television (CCTV), Southern Media Corporation (SMC) and China Radio International (CRI). iSuppli, a market research agency, said that by the end of December, 2007, the number of IPTV users in Chinese mainland had reached 846 thousand. There are 151.18 million cable TV users in China, 362 million families possess TV sets by 2007 with an annual growth rate of 10%; China have 2058 TV channels with a coverage rate of 96.6%; By the end of 2007, China had 26.16 million users of cable digital TV, accounting for 24% of all cable TV users. In China 10 million hours of TV programs are broadcast each year, and the total of TV audiences above 4 ages old reached 1,199 million, accounting for 95.84% of all people above 4 years old; Average daily TV-watching time per person was 176 minutes.

5 METADATA AND DATABASE

Metadata is data about data. Libraries' integration with telecom and network, adopting a new service mode, needs to firstly classify online resources so as to make readers know the contents, features, and functions of information through describing metadata. No matter how complicated metadata may be, it is made up of the following parts of information resources: 1) title, abstract, key words, 2) publisher, website address, and metadata's maintenance information. Metadata is different in Chinese libraries from that in other countries' library. For example, we use CNMARC (Machine Readable Catalog Format in China) to record Chinese books and reference materials, and use

USMARC or UNMARC to record Western language books. This method is being used in computer catalogue. Most of the libraries in the world use Dublin Core Elements Set (DCES) to describe network resources, DCES include 15 elements: Title, Creator, Subject and Keywords, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage and Rights (Qiu Dongjiang, 2006). Metadata retrieval service is the core content of information resources' catalogue service. The US developed international standard for network retrieval. It's called "Information Retrieval: Application Service Definition and Protocol Specification for Open System Interconnection" (ANSI/ NISOZ39.50-1995), and it was accepted by International Standardization Organization (ISO) and named ISO23950 in 1991, and was used in libraries extensively. Z39.50 includes 8 types of operations: 1. Init (Initialization), 2. Search 3. Select 4. Delete 5. Scan 6. Sort 7. Resource report 8.Extended services.

Databases are datasets that organize information resources according to certain technology standards and deposit them in computer storage device. ISO/DIS5127 defines a database as "a kind of datasets that are made up of at least one kind of document, and can meet a certain special objective or the need of data processing system." Databases include two kinds: Reference Database, Source Database。Reference Database includes Bibliographic Database and Referral Database; Source Database includes Numeric Database, Textual-numeric Database, Full-text Database, Terminological Bank, Graphics Database and Multimedia Database. Readers retrieve databases mainly through three major systems: optical disc (CD-ROM Jukebox, CD-ROM Tower and CD-ROM Mirror image), on-line service (DIALOG, MEDLARS, STN, BRS and OCLC) and Web-database (Directory search engine, Robot search engine and Meta search engine).

6 SEMANTIC EXPANSION OF DATA

The explanation of "data" in <The contemporary Chinese Dictionary>: 数据 shùjù, data; usu Used as basis for calculations in statistics, planning, scientific research, technical design, etc. While in <New multi-purpose English-Chinese Dictionary>: data ['deitə] n pl. (sing. datum) facts; things certainly known. datum ['deitəm][the same as]information, facts; figures. Here "data", "information", "facts", and "figures" are synonyms. In two dictionaries above: 信息 xìn xī, information; news; message (of information theory) information transmitted by signals with the content not previously known to the recipient; information [ˌinfə'meiʃən] n. something told, news [synonym] report, message knowledge got by search, study, etc. facts learnt [synonym] data, facts, knowledge, documents, materials (law) a complaint or charge in a law court [synonym] denunciation, accusation. 事实 shì shí, fact: Facts are stronger than rhetoric; fact [fækt] n. something that has actual existence or an event that has actually happened or is happening; something true [synonym] happening information regarded as being true and as having reality the truth. In English, "information" was used to indicate Chinese word "信息". This word has several meanings in English, (1) news (2) intelligence (3)knowledge (4)report (5)speech (6)lecture (7)notice (8)service platform (9)inform, tell (10) [computer] information. It is clear that "information" is a word of wide range. The word "information" has the most extensive scope among the three synonyms of "data", for it almost covers all contents of books and internet resources, but in practice, sometimes they are synonymous, sometimes they are different, so the three words interpenetrates each other. When we classify hard copy literature with CLC (Chinese Library Classification), LCC(Library of Congress Classification) and DDC(Dewey Decimal Classification), and classify network resources with DCES(Dublin Metadata Core Element Set), data is an important tool for information management, the semantic meaning of data is being extended from only numerical value and figures to words and expressions relating to linguistic characters, symbol system, design and images, diagrams, sound, video and etc. This point is especially outstanding in the operation of library database management system. The actual integration of data and information is a linguistic appearance of the integration of library, telecom, and internet. The semantic extension of data will open up vast space for scientific data research.

7 REFERENCES

CNNIC (China Internet Network Information Center). (2008) the Twenty-first Report on Internet Development Status in China. Retrieved Oct.30, 2008 from the World Wide Web: http://www.cnnic.cn/uploadfiles/doc/2008/1/17/104126.doc

Chen Xin. (2008) Observation of new trends of publication from US digital publishing status. *Wen Hui Bao*. Jan. 20(1)

NBSC (National Bureau of Statistics of China).(2008) *Statistical Communiqué of the People's Republic of China on the 2007 National Economic and Social Development*. Retrieved Oct. 29 2008 from the World Wide Web: http://www.stats.gov.cn/tjgb/ndtjgb/qgndtjgb/t20080228_402464933.htm

Qiu Dongjiang. (2006) *The new library and intelligence science dictionary*. Beijing, China: Science and Technology Literature Publishing House

Wang Dasheng. & Liu Bin. (2005) Information management and culture productivity. *Discovery* Supplementary Issue (12), pp 7-10

Wang Dasheng. (2006) The integration of Library, network and telecom. *Journal of Information*. (First volume), pp 109-111

Wang Xianchong. (1986) The theory and use of electromagnetic-field. Beijing, China: Science Press.

Wang Xuewen. (2002) The analysis about media structure and market quota of China. Retrieved Oct. 15, 2006 from the World Wide Web: http://www .gsei.com.cn/aboutgansu/ziliao/gqbgk/200406/504166.htm

Yan Xiaohong. (2007) The status, questions, and countermeasures of Chinese digital publishing industry. Retrieved Oct. 15, 2006 from the World Wide Web: http://chinabook.gapp.gov.cn/Web/ArticleShow.aspx?artid=066727&cateid=A18

Zhang Le. (2007) The total storage of e-books of Universal Library exceeded 1.5 million copies. Retrieved Sept. 15, 2008 from the World Wide Web: http://news.xinhuanet.com/newscenter/2007-11/28/content_7162908.htm

Zhan Furui. (2008) The library industry in China: an unordinary year of 2007. *China Culture Daily*. Jan.20 (1)