World Library and Information Congress:
71th IFLA General Conference and Council

"Libraries - A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

Conference Programme:
http://www.ifla.org/IV/ifla71/Programme.htm

August 4, 2005

Code Number: 170-E
Meeting: 143 Asia and Oceania

Sustainable Digital Library Development for Scientific Communities

Dr Zhang Xiaolin
Library of Chinese Academy of Sciences, Beijing, China
zhangxl@mail.las.ac.cn

Abstract:

The paper analyses the digital library services for China's scientific communities, explores the challenges toward sustainable digital libraries, presents a development strategy for digital libraries using Chinese Academy of Sciences as an example, and provides a layered design framework for a digital library in an e-science environment.

Keywords: digital library, sustainable development, Chinese Academy of Sciences

1. Trends of Chinese Digital Library Development for Science & Technology

Generally speaking, digital library development in China can be grouped into local DL projects, national specialized DL systems, and national DL programs. While local DL projects have produced rich digital cultural materials from local libraries that can be accessed through network, it is the national ones that make possible the wide and fast increasing availability of digital information resources, especially for scientific research and educational institutes.

Since late 1990s several national DL efforts have been made that gradually brought about a national digital and networked information infrastructure in place to support the nation's scientific research and education. Here are a few examples.

Chinese Science Digital Library (CSDL)[1], initiated in 2001 by Chinese Academy of Sciences (CAS) to develop a digital information service platform to serve its 91 research institutes across the country. At the current stage, CSDL provides CAS researchers with full text STM journals, conference proceedings, theses and dissertations (ETDs), patents, reference books, and e-books. For e-journals alone, CSDL now covers 4000 core western STM journals and 10000 Chinese ones. At the same time CSDL organizes a wide range of networked services including union catalogs, federated database search, document delivery, digital reference, MyLibrary customization, and remote authentication.

Chinese Academic Library and Information System (CALIS)[2], first funded in 1998 by the Ministry of Education (MoE) to organize a digital library consortium among academic libraries. CALIS membership has grown very fast to 500 in 2005 with most of them joining together under CALIS sponsorship to collectively negotiate acquisitions of e-journals, e-books, ETDs, etc., and many of them also taking part in building specialized subject gateways or digitization of special educational and cultural materials. With the help of CALIS, readers at most Chinese universities can now access to far more materials than their print collections can provide. Again taking journals as an example, the average number of print foreign titles stands less than 1000 but most universities can access to more than 5000 foreign e-journals while many can cover more than 10000 titles.

The National Science and Technology Library (NSTL)[3], is established in 2000 by the Ministry of Science and Technology (MoST). NSTL is a virtual library consisting of 7 special libraries from basic sciences, agricultural sciences, medial sciences, and engineering, to form a national reserve library and a supply center for scientific information. NSTL supports its member libraries with acquisition of additional foreign journals, production of a central abstract database for all 13000 foreign journals in its members' collections, and supply of requested documents with only nominal fees and fast turn-around time. Since NSTL is open to all in the country as a reliable and timely service, its entire collaborative collection of foreign and Chinese materials serves as a back-up system for all libraries, even including special digital library systems such as CSDL and CALIS, whom now will concentrate on building core e-resources that can be retrieved immediately by their clienteles while relying on NSTL to provide additional coverage.

With systems like NSTL, CSDL, and CALIS, the information environment is changing very rapidly for Chinese scientific communities, especially those in advanced research institutes and universities:

(1) The dominant information resources for scientific users are now digital;

(2) The default and preferred way of accessing information for scientific users are now through network;

(3) Most information services, even including reference services, are increasingly provided by collaborative effort among libraries (and of course through network);

(4) Most physical libraries are increasingly integrated into virtual digital information systems which are now the first entry point for most scientific users.

While this trend has created a far more effective information space for scientific users and has provided libraries with great accessibility power, it also presents a revolutionary transformation of the information environment and of the ways libraries define themselves and operate their services. It is no longer "business as usual". Clear analysis of the challenges brought by the changes is a critical prelude to any effort trying to sustain digital the development for libraries, digital or otherwise.

2. Challenges to Sustain and Develop Library Services

Challenges come from two major directions. The first is with effectively managing and sustaining the digital library itself to serve traditional information needs, and the second goes beyond traditional sense of library services to support new user needs and behaviors in the new e-science environment.

When digital libraries become the way of life for users as well as for libraries[4], we are facing with many new tasks that require drastically different knowledge, methods, and mechanisms. Just mention a few:

(1) Resource development becomes primarily a consortium effort so that any one library has to plan and work together with other (often different) libraries based on systematic analysis of one's needs and supply chain structures.

(2) Resource structures of libraries are changing increasingly with the steadily transformation from the print plus model to e-plus model and to e-only model, which asks for new strategies, new supply-chain structures, new funding models, new organizational procedures, and new coordinative infrastructures for print and e-copies.

(3) Digital preservation becomes critical and regular operations for a sustainable library service, which requires new system infrastructure, financial support, technological and administrative expertise, and commitment to long term planning and management.

(4) Preservation of imported information resources manifests as another challenge for any country where much scientific information comes from foreign publishers and most acquisitions of databases are only for the right to access. Securing perpetual copies for one's scientific communities is of a national importance and a network of shared responsibilities among distributed trustable archives is urgently needed.

(5) The first line of services switches from physical reading rooms to websites and the main stream of services transfers from passive reading and circulation to active and interactive personalized services such as digital reference, selected dissemination of information, information analysis, specialized portals, virtual reading rooms, and information literacy. Collection-oriented workflows and library-centered service structures, though a century old, are no longer adequate, and new user-centered and service-oriented restructuring of libraries are called upon.

However, addressing only challenges like these is not adequate for libraries to sustain their development in the face of the drastic changes in the very environment where information is created and used. Fundamental changes lie ahead which calls inevitably for new models of digital libraries:

(1) Science is increasingly operating in a digital fashion. E-science or cyber-infrastructure[5] creates a networked infrastructure rich in data grids, computing grids, digital libraries, and collaboratories. Here all kinds of scientific objects, including people, programs, facilities, data and documents, procedures and workflows, and even policies and strategies, can be and will be digitally represented, accessed, interconnected, and invoked through networked interaction. This calls for a new definition of information resources, information organization, and information service integration.

(2) Science is increasingly based on interactive virtual knowledge communities[6] where networked information becomes interactive research tools, online collaboration acts as the organizing mechanism, and virtual organizations dynamically supports knowledge discovery and exchange. Effective services far beyond simple search and delivery of known items are demanded to help users mine knowledge and knowledge relations among seas of various scientific objects, and do so proactively along research workflows and among research interactions. A service system diffusing into and interacting with users' knowledge processes are essential for success in such a situation..

(3) Scholarly communications is taken new turns when forces like Google Scholar/Print[7], open access materials[8], and institutional repositories[9] are creating a new information supply chain. Access to information is no longer intermediated by and channeled through a library, "library services" can be more effectively provided by open or commercial systems, a distributed, producer-driven, value-enriched, and competitive market is here to stay, and providing access to information alone is no longer enough for a sustainable future, as predicted by PEW study that any organization relying on intermediary services will be fundamentally changed within near future[10].

(4) Digital libraries as they are today will be hard pressed. In many ways, most digital libraries are modeled after traditional libraries. While recognizing the merit of this, one has to realize the very nature and limitation of such digital libraries being resource-centric, document-centric, and library-centric. Facing the new environment, digital libraries have to contribute more directly and uniquely to the core knowledge processes of users than what can be done by other ISP/ICPs, have to organize and present digital content in context-sensitive and knowledge-aware manners, and have to extend vision of information activities and information services into the totality of users' problem-solving processes and users' information environment. Actually, digital libraries need to be re-designed as an integral part of users' knowledge work platforms instead of stand-alones or add-on extras[11].

3. Remodeling the strategy for digital library development

New models of information and knowledge services have to be found or invented. Libraries can either lead the way or being marginalized on the way. Library of Chinese Academy of Sciences (LCAS) takes these challenges as opportunities to develop its digital library strategy by envisioning a process to transform digital libraries into an e-knowledge infrastructure where:

(1) Digital objects in the research space are described, organized, interlinked and integrated as an open knowledge grid.

(2) Networked services are provided in a personalizable and user-based way to retrieve, organize, link, visualize, re-purpose and re-distribute information on digital objects in the research space.

(3) Knowledge processing mechanisms are built within user information environment and embedded into user information processes to support knowledge discovery, knowledge organization and knowledge management.

(4) Collaborative mechanisms are developed for librarians to work together with researchers to develop the above structures and services, and to provide further value-added and personalized services.

Aiming to build a digital library diffusing into research environment, LCAS investigated development trends in e-science and digital libraries and surveyed CAS users for their current and future needs. Based on a new understand of the essence of digital libraries as a live knowledge system incorporating knowledge content, context, and communities[12], the following framework of CAS Digital Library is outlined.

(1) General Layered Framework (Fig. 1). It consists of three major layers: a common information layer, a discipline-based scholarly communication layer, and an institution-based knowledge layer. An information analysis module is provided to aid knowledge discovery, an ontological integration service is planed to provide the semantic glue to integrate resources and services at various layers, and a workflow management service module is envisioned to connect information processes between various layers.

(2) The common information layer is made of two parts: firstly, a supporting mechanism for a distributed CAS information resource system that provides basic development, organization, access, preservation capabilities for key information resources needed by CAS users. Secondly, an integrated service platform is built upon the resource system layer but extends to integrate resources from NSTL, CALIS, and other digital libraries, to provide a logically coherent and integrated information services layer no matter where the resource is. This is based on today's CSDL which is already in place.

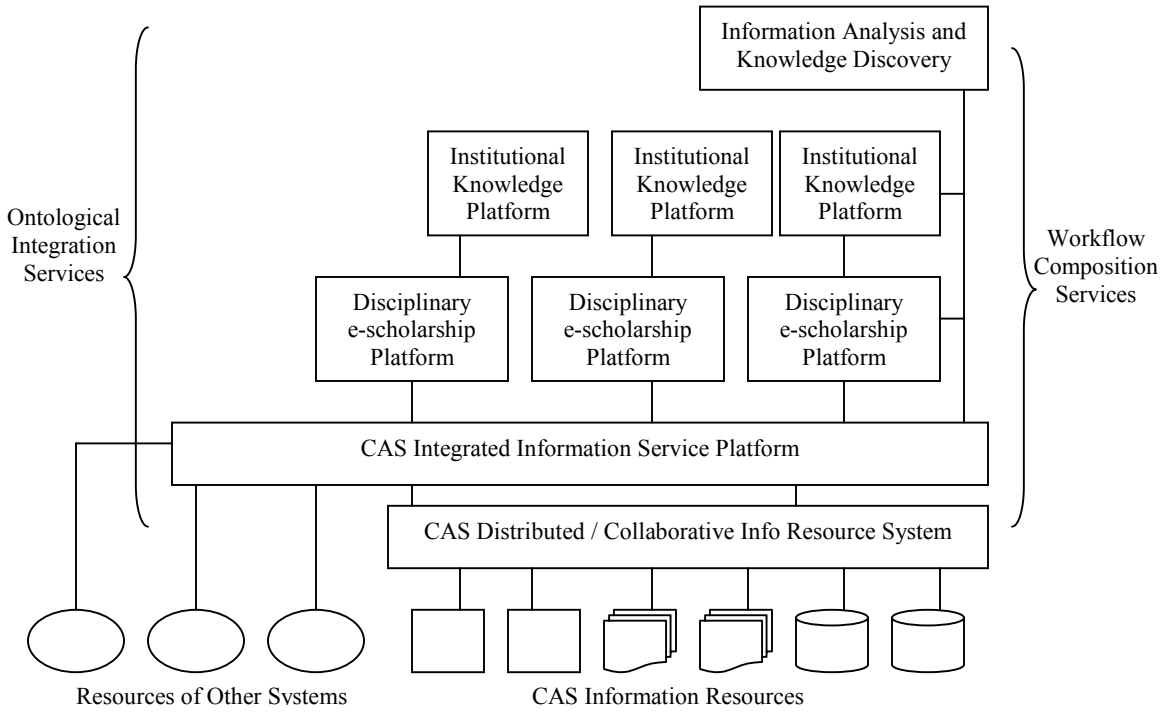

Fig.1 CAS Digital Library Framework

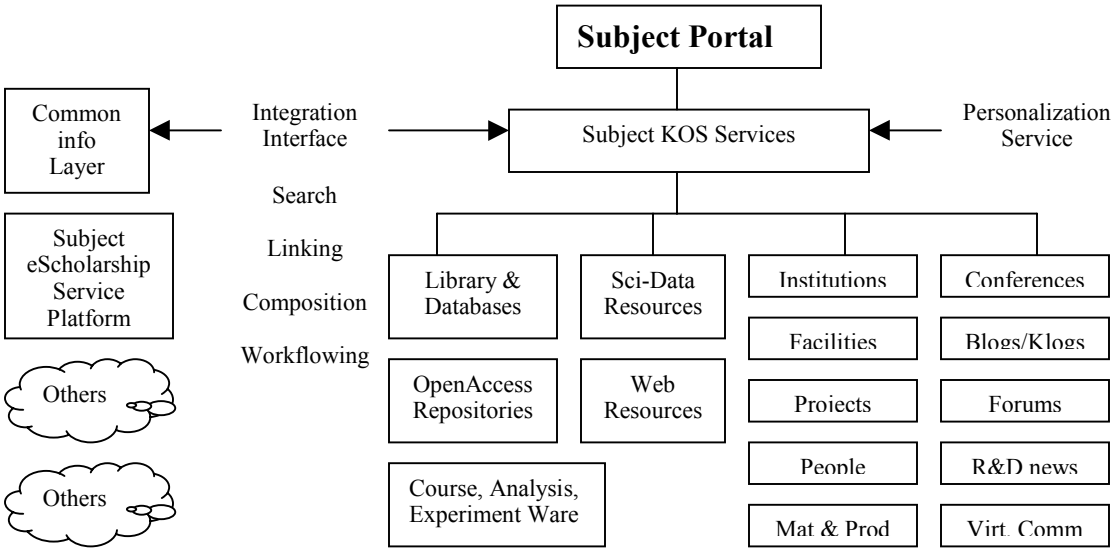

Fig.2 Disciplinary eScholarship Platform

(3) Disciplinary e-Scholarship Platform Layer (Fig. 2) is a series of disciplinary e-scholarship platforms. "Disciplinary" may mean disciplines, subjects, problem-areas, or topic-areas, but the emphasis here is to build around the scholarly communications process to support the whole range of user scholarly communications activities. So this layer adds a whole range of informational resources (scientific data, open access repositories, etc.), research resources (analytic tools, institutions, facilities, projects,

people, materials and products), and communicational resources (conferences, forums, blog/klogs, R&D news services, and virtual communities).

(4) Institutional Knowledge Layer (Fig. 3) is a series of institutional knowledge platforms that serve individual CAS research institutes, laboratories, or project teams. The emphasis here is on knowledge management from a community and context point of view. It also integrates information resources, research resources, and communications resources, but further localized according to the needs and context of a problem-solving community. Then a series of knowledge production entities are organized into the platform, such as the institution's departments, projects, teams, people, and facilities. A system of living knowledge entities is further organized into the platform, such as plans, activities, processes, news, and results. The word "living" is important, for their dynamic changes, for linking their changes, and for linking them with other entities. To promote active user knowledge processing and utilization, a number of "Wares" are provided. An institutional repository stands at the heart to organize, manage, and preserve institutional knowledge properties that may come from the living entities within or from sources outside.

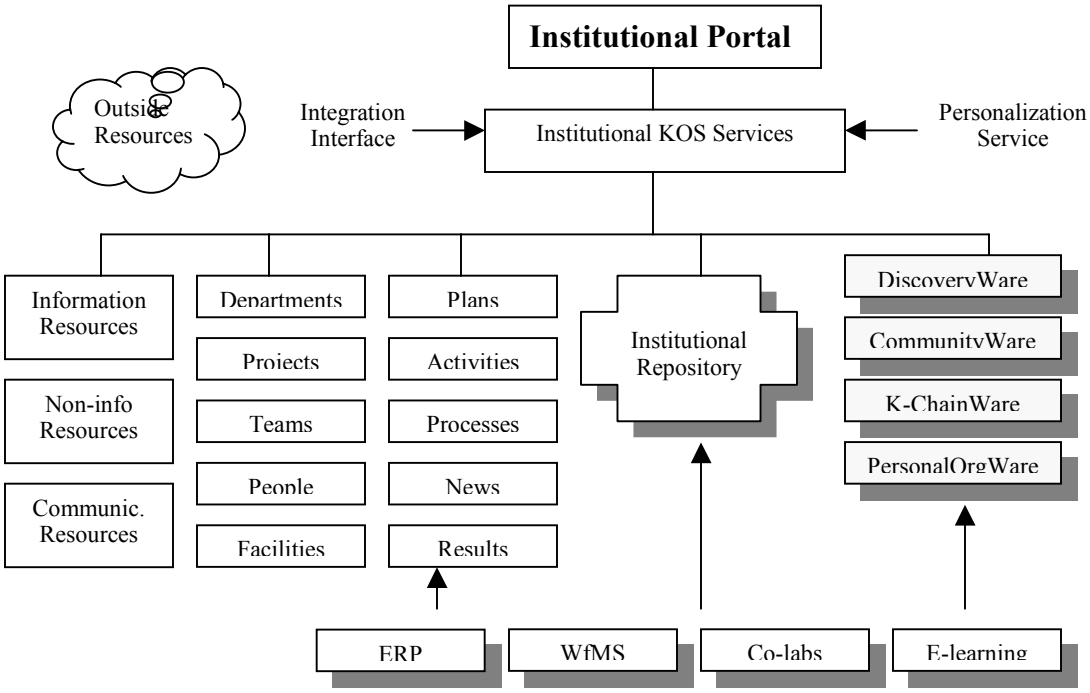

Fig.3 Institutional Knowledge Platform

This is indeed an ambitious design that can only be implemented phase-wise. But what important is the vision and the approach for digital library development from user knowledge processes. It is the author's believe that the viability and sustainability of digital libraries lie in their diffusing into and support for user knowledge processes. The design outlined here is an experimental step toward this goal.

References:

[1] Chinese Science Digital Library http://www.csdl.ac.cn/
[2] Chinese Academic Library and Information system. http://www.csdl.edu.cn/

[3] National Science and Technology Library http://www.nstl.gov.cn/
[4] 2003 OCLC Environmental Scan: Pattern Recognition.
http://www.oclc.org/membership/escan/toc.htm
[5] Atkins, D. E. et al. Revolutionizing Science and Engineering Through Cyberinfrastructure. National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure, January 2003.
[6] Atkins, D. E. Research Libraries and Digital Library Research for Cyberinfrastructure-enabled Knowledge Communities. In The Proceedings of Digital Library-Advances the Efficiency of Knowledge Utilization, ed. Furui Zhan, Beijing: Scientific Documentation Press, 2004.9
[10] PEW Internet & American Life Project: The Future of the Internet, January 2005.1
[11] Lougee, W. P. Diffuse Libraries: Emergent Roles for the Research Library in the Digital Age. Washington, D. C. Council on Library and Information Resources, August, 2002.
[12] Norris, D., Mason, J. and Lefrere, P. Transforming e-Knowledge.
http://www.transformingeknowledge.info/index.html