

Digital Libraries in Japan and at ULIS

Shigeo Sugimoto

University of Library and Information Science (ULIS)

Tsukuba, Ibaraki 305-8550, Japan

sugimoto@ulis.ac.jp

Abstract

Digital Library is widely recognized as a key issue for building knowledge and information infrastructure in Japan as well as in other countries. There are a numerous activities on digital library in Japan. The aim of this paper is to overview the research and development of digital libraries in Japan and at ULIS. This paper firstly overviews the research and development activities of digital library in Japan, which include scholarly information oriented digital libraries at the National Center for Information Systems (NACSIS) and university libraries, National Diet Library (NDL), Ministry of International Trade and Industries (MITI), and so forth. This paper secondly presents the digital library service at the university library of ULIS whose important function is a subject gateway for library and LIS related resources. Thirdly, this paper shows other activities on digital library at ULIS.

Keywords

Research and Development Projects of Digital Libraries, Information Technologies, Scholarly Information, Japan, University of Library and Information Science, Subject Gateway, Metadata, Dublin Core, MHTML

1. Introduction

In Japan as well as in other countries, digital library is widely recognized as a key issue for building knowledge and information infrastructure. There are a numerous

activities on digital library in Japan. Those activities range in size from small-scale to nation-wide, and in type from humanities oriented digitization to information technology research.

Research and development of digital libraries can be classified into three categories,

- Collection development centered,
- Information technology centered, and
- User and social issues centered.

Collection development activities mostly have their basis in libraries or library related organizations. In Japan, the important players are university libraries and national institutions such as National Center for Science Information Systems (NACSIS) and National Diet Library (NDL). Information technology centered activities are mostly found in computer science laboratories in universities and industries. The R&D area of information technologies for digital library covers quite broad range of technologies, i.e., IR, OCR, document processing/understanding, natural language processing, user interface and visualization, intellectual property handling, electronic trading, and so forth. In Japan, there are no large IT centered R&D project except the Next Generation Digital Library Technology project by IPA and JIPDEC [1]. User and social issues are quite important for making digital libraries really usable. However, from the author's viewpoint, they are not so dominant as the other issues. Development of digital libraries and digital collections will promote researches in this aspect.

This paper surveys the activities on digital libraries in Japan and at ULIS. It describes, for example, R&D of information technologies, digital collection development and digital library services at libraries. Social issues and user issues, such as intellectual property, ethical and privacy issues, are out of the scope of this paper.

2. Digital Libraries in Japan – an overview

Digital collection development of scholarly information is a central topic for the Japanese library community. University libraries, NACSIS, NDL are key players for digital collection development. The Ministry of Education, Science, Sports and Culture (Monbusho) has funded six university libraries to build and operate digital library services. NACSIS has started the commercial digital library service which provide journal articles published by Japanese academic societies and institutions. NDL is working on their new digital library services which will be one of important functions at their new library Kansai-kan. National Museum of Ethnology is working on the Global Digital Museum with British Museum and IBM [2]. Agriculture, Forestry and Fisheries Research Information Center (AFFRIC) has their Network Library System which provides various information resources including full-text articles published by the research institutes of the Ministry of Agriculture, Forestry and Fisheries, and information about WWW resources related to their subject areas [3].

Digital library is widely recognized as an important application area of information technologies. IT research in the related areas is active, e.g. filtering and dissemination to support information access, multi-lingual information retrieval and access, and so forth. Information technology development for the next generation digital library system at JIPDEC is a large IT oriented project. However, there is no other significantly large scale research. Information-technology Promotion Agency (IPA) had experimentally developed a large scale collection of digital images. This digitization project called Electronic Library Pilot Project was

cooperatively performed with NDL and digitized about 10 Mpages [4]. BBCC, which is a non-profit organization cooperatively working with Ministry of Post and Telecommunication to promote broadband telecommunication network, has hosted cooperative experimental research of digital libraries [5].

Public libraries are seeking new ways to provide digital information and Internet access to their users. Aozora Bunko, which is a voluntary activity, provides collection of electronic text of Japanese literatures [6].

3. Ministry of Education, Science, Sports and Culture (Monbusho)

3.1 Overview

Monbusho has been promoting development of digital libraries and information environment at university and college libraries, museums and related institutions. These institutions are important players in digital libraries, since they have a large collection of scholarly materials suitable to digitization and a large number of users, i.e., researchers and students [7].

In 1996, Nara Institute of Science and Technology (NAIST) launched the first digital library service among the Japanese national universities. In July 1996, the Science Council recommended to Monbusho that university libraries need to enhance their functions and digital information resources. In 1998, Kyoto and Tsukuba universities started digital library services at their university libraries. In 1999, three universities, Tokyo Institute of Technology, Kobe Univ., and ULIS, started to build digital libraries. The activities at university libraries are described in section 3.3.

National Center for Science Information Systems (NACSIS), which maintains a national academic backbone network and databases including union catalog for university libraries, is an information hub for Japanese universities. NACSIS has been working on a digital library project named

NACSIS-ELS [8][9] and has started the commercial service of the NACSIS-ELS from 1999. NACSIS-ELS provides a digital collection of journals published by Japanese academic societies of broader subject range, from the humanities to engineering and medicine. The activities of NACSIS are surveyed in the following section.

Museums and related institutions such as National Museum of Ethnology and National Institute of Japanese Literature are working on building digital collections and information access environments.

3.2 National Center for Science Information Systems (NACSIS)

As described above, NACSIS is a hub of scholarly information for the Japanese academic community. As of August 1999, NACSIS-ELS provides more than 260 titles published by 85 academic societies and institutions¹.

NACSIS Online Journal project (NACSIS-OLJ) provides an integrated software environment for authors, editors, and publishes to publish scholarly journals online [10]. It has potential to promote online journal publishing not only by large academic societies and institutions but also small ones which cannot afford technologies for online publishing.

NACSIS Test Collection for IR Systems (NTCIR) is a project to build a large test collection of documents to promote research on IR including Japanese-English cross-lingual IR. A competition-type workshop has been held using the NACSIS Test Collection 1[11].

3.3 University Libraries

University libraries, especially national university libraries have been playing a very important role for development of digital libraries in Japan as well as in other countries.

¹ The number of titles include titles already available and those that will be made available.

Central part of the digital library development at university libraries is collection development. As universities differ in their sizes, histories and the subject areas, the features of the digital library reflect the difference. The following paragraphs describe the features of the digital libraries at university libraries.

(1) Nara Institute of Science and Technology (NAIST)

NAIST is a new graduate school specialized in computer and information sciences, bio-science, and material science. The digital library at the NAIST library provides papers included in journals and magazines, theses and dissertations published in NAIST, video and multimedia materials[12][13]. The research and development section is building a video collection of lectures at classrooms as a part of the Internet University project which is a cross-university research project with other participating universities [14]

(2) Kyoto University

An experimental digital library called Ariadne had been developed in Kyoto University prior to the development of the operational system [15]. Ariadne had advanced features such as machine translation between Japanese and English, retrieval function enhanced with document structure, and so forth. In 1998, Kyoto University Library started the development of their operational digital library [16]. Kyoto University digital library provides a digital collection of cultural heritage materials and a collection of publications produced by university members. From a technological viewpoint, this digital library has advanced retrieval functions; for example, contextual search based on a hierarchical structure of a table of contents of a document, search by terms automatically translated between Japanese and English, and so forth.

(3) Tsukuba University

Tsukuba University Library started their digital library service in 1998. The central function of this service is to make materials published in the university accessible via the Internet. Their digital collection includes

dissertations and other material on campus. It also includes historical materials digitized in high resolution format, i.e. 1100 dpi.

(4) Tokyo Institute of Technology (TITech)

TITech is a large institution specialized in science and technology. The TITech Library has been appointed as a hub library to collect science and technology (S&T) publications, e.g. journals, magazines and proceedings. The digital library at TITech library is oriented to S&T materials, i.e., journals and proceedings in S&T areas, dissertations and technical reports published on campus. A subject gateway service for S&T resources is also included as their central function.

(5) Kobe University

Kobe University Library is planning to provide a digital collection of Kobe Quake materials, economics resources, and other rare historical materials. The most significant feature of this service would be the Kobe Quake collection. The collection will include various kinds of materials related to the quake, which include books, magazines, brochures and pamphlets, posters, photographs, maps, videos, and so forth.

(6) University of Library and Information Science (ULIS)

The Digital Library service at the ULIS Library (ULIS-DL) started in February 1999. The principal purpose of ULIS-DL is to build up a subject gateway for LIS and libraries, which provides a collection of metadata of information resources in LIS and for libraries. The metadata is created based on Dublin Core. ULIS-DL is also designed to provide a collection of digital resources, e.g. scholarly materials published on campus, lecture notes and course materials, databases, and so forth. ULIS-DL is described in more detail in section 6.1.

(7) Other University Libraries

University libraries are making large efforts to provide information resources via the Internet for on- and off-campus users. For examples, University of Tokyo Library provides metadata of academic resources, Hokkaido University Library provides Northern Studies Collection,

National university libraries in Kyushu are working together to promote digital library services as the regionally organized cooperative services, and so forth. Thus, university libraries are promoting digital library services and the amount of digital resources are increasing. In parallel with the promotion of those digital library services, it has been recognized as an important issue to promote collaboration among the libraries and to share information, especially metadata. Some meetings have been held to exchange information, know-hows, and ideas for future collaboration.

4. National Diet Library (NDL)

Like national libraries of other countries, NDL is an important player in the Japanese digital library community and for international cooperation. For example,

- Digital library will be one of the key functions for their new library Kansai-Kan [17][18],
- NDL has been working on the digital library at the International Library of Children's Literature, and
- NDL is a core member of the G8 digital library program.

Kansai-kan is a new branch library planned to open in 2002. It is located in the Kansai Science City near Osaka and Kyoto, which is 500 km west from the main library of NDL in Tokyo. Digital library is one of the core functions of Kansai-kan to cope with decentralization of holdings, staffs and users and to enhance service quality for domestic and international users².

NDL will found the International Library of Children's Literature in January 2000 as a branch library of NDL [20]. The library will collect domestic and foreign literature for children and related materials. It will have digital archive and digital museum for the children's materials, which will be made accessible via networks. NDL has carried out a pilot project for the children's library in 1995

² Documents on digital libraries published by NDL are accessible on WWW [19].

and 96. The pilot project included the bibliographic data of about 40,000 volumes of juvenile books published before 1976, digitized data of about 9000 volumes which include books published before 1950 and picture books published before 1955. Experimental systems were also included in the pilot; they are multilingual picture books for the Internet with oral narration, and an image-based navigation system to help children find and select books.

NDL's digital library projects has potential to promote not only development of information technologies and digital collection but also development of social systems such as legal deposit of electronically published materials, intellectual property and economic models for libraries and global user services.

5. Ministry of International Trade and Industries (MITI)

The Information-technology Promotion Agency (IPA), which is conducted by MITI, had a cooperative project with NDL on a pilot digital library which includes a large collection of digitized materials and a union catalog for public libraries. The digital collection includes about 10 million images of materials ranging from old rare manuscripts to books published during the World War II and comics published in 60's and 70's. In 1998, development of the digital collection has been completed.

IPA has an R&D project for new generation information technology for digital libraries. The project is being carried out by the Japan Information Processing Development Center (JIPDEC). JIPDEC has been working with information and computer companies to develop new technologies for DL and a prototype digital library system which is an integration of the new and existing technologies. The purpose of this project is to find an software architecture and application technology for a large scale digital library system, e.g. NDL's digital library. The R&D topics include basic system architecture, agent-based software architecture, advanced information retrieval technology, digital document technologies, human interface,

intellectual property right management, and system integration.

6. Digital Library and Related Activities at ULIS

6.1 ULIS Digital Library (ULIS-DL)

ULIS-DL is designed to provide a collection of metadata, a collection of primary resources, a system environment to create digital materials in various kind of information media, and information access environment in reading rooms [21].

As described previously, the subject gateway service is the primary function of ULIS-DL. We collect resources and create metadata records of the resources. The major resources are WWW documents contained in the WWW sites of libraries and LIS-related institutions. The metadata element set, called ULIS Core (UC), is defined based on the 15 Simple Dublin Core elements with small extensions. The extensions are character code and country of publication as primary elements, and pronunciation information as a sub-element of every element. The metadata records are basically created for every document object collected by a crawler.

From a technological viewpoint, the gateway requires an integrated support environment for producing and maintaining the metadata. In ULIS-DL, a resource collected by the crawler is given to a software which automatically extracts information from the resource and creates a metadata record based on UC. The automatically created metadata record, which is called metafile, is given to a human cataloger. The cataloger creates a metadata record for the source resource viewing the metafile and the resource. The created metadata record is added to the metadata database after approval by a manager. The system has a set of support tools for maintaining the metadata records, e.g., to check if a URL written in a metadata record is valid, to check if a source resource is updated, and so forth.

Language is obviously a crucial issue for this

system since WWW resources of our interests are published in various languages. We mainly collect resources written in Japanese and English. Most of the metadata is written in Japanese. The user interfaces for retrieval have machine translation functions between English and Japanese for users who do not understand Japanese.

6.2 Promotion of Information Exchange on Digital Libraries

Digital Library Network (DLnet) was launched in ULIS in 1994 as an activity to promote information exchange and knowledge sharing among researchers, developers, practitioners and users who have interests in digital libraries, and to build up a forum to discuss digital libraries and related matters [22]. The series of Workshop on Digital Libraries (DLW) have been held mostly at ULIS based on the DLnet. 15 workshops have been held since August 1994. The 12th and 15th DLWs were held at Asian Institute of Technology (AIT) in Thailand and NAIST, respectively. The workshop language of all DLWs except the 12th DLW was Japanese. The 12th DLW was held as a part of the International Joint Workshop on Digital Libraries 1998 which was organized by AIT, Kasetsart University, and the Thai Library Association. On the other hand, the international symposia on digital libraries, ISDL'95, '97, and '99 have been hosted at ULIS. At these symposia, we had speakers from North America, Europe, Asia-Pacific and Japan. Most of the papers published at these conferences are accessible via WWW.

6.3 Multilingual Browser for WWW Documents (MHTML) and Japanese Oldtales Collection in Multiple Languages

A browser technology to display multilingual HTML texts on off-the-shelf browsers has been developed in ULIS [23][24][25][26]. Users can display foreign language texts on their own browsers without installing fonts for the foreign languages. The technology, which is called MHTML, has been enhanced to provide multilingual text input functions, e.g., a Japanese text input function from a client

machine that has no built-in Japanese text input function. Since the MHTML browser has been implemented as a Java applet, the browser is platform independent and users can ubiquitously access foreign or multilingual texts via MHTML running on off-the-shelf browsers.

A collection of Japanese old folktales written in multiple languages has been built using the MHTML technology[27][28]. As of August 1999, 11 tales in 11 languages, the collection includes 10 Japanese popular old tales, eg., Momotaro, Issunboshi and Cracking Mountain, and one tale written by a volunteer author. The languages include European languages, such as English, French, German, and Spanish, and Asian languages, such as Chinese, Korean, and Japanese. The old tales were chosen from well-known tales and originally written by volunteer authors, and then translated by volunteer translators. This type of activity could not be accomplished without the global information access and international collaboration via the Internet.

Dublin Core Metadata Element Set is a metadata element set defined for resource discovery on the Internet. Metadata description in multiple languages is widely recognized as an important issue for global resource discovery in the Dublin Core community. As a first step, the working group on multiple languages (DCML) is developing a registry of reference descriptions in multiple languages [29]. The MHTML technology is used to display the reference description in multiple languages on off-the-shelf browsers [30].

7. Summary

We can find many activities related to digital library other than those described in this survey. The activities include R&D of information technologies for digital libraries, digital collection development, and so forth. We can also find collaboration effort among libraries to share resources and ideas. However, collaborative efforts by IT researchers and libraries are rarely found. The author believes collaboration between

researchers and practitioners of IT and library is crucial to promote digital library.

ULIS-DL is still in the beginning phase, but the following issues are crucial for further development,

- evaluation of quality and usability of metadata,
- research on evolution of metadata description scheme, i.e. Resource Description Framework and sub-elements for qualified description of metadata, and
- collaboration with other libraries and subject gateways in Japan and in the world.

Acknowledgments

I would like to express sincere thanks to the following people. Professor Koichi Tabata founded the series of Workshop on Digital Libraries and led the MHTML technology development. Tetsuo Sakaguchi, Norihiko Uda, Mitsuharu Nagamori, Takehisa Fujita, Akira Maeda, Myriam Dartois who have been and are important members to support the conferences and MHTML. Professors Jun Adachi, Masamitsu Negishi, Keizo Oyama and other members of NACSIS helped us start DLnet and organize the conferences. Professors Masakazu Imai of NAIST, Vilas Wuwongse of AIT, Asanee Kawtrakul of Kasetsart University, Thomas Baker of GMD played important roles for domestic and international cooperation. Hideho Ohba, Toshiaki Yokoyama and Hiroshi Hiraoka of ULIS Library and Masayuki Kawase of Kyushu University Library have made great efforts to realize and manage the ULIS-DL. I would like to express my thanks to the executives of ULIS for their leadership to realize the ULIS-DL and help to other activities. I would like to thank JIPDEC and NDL for their cooperation.

References

- [1] Mukaiyama, H., Technical Aspect of Next Generation Digital Library Project, Proc. of ISDL'97, pp.72-79, November 1997
- [2] Junichi Takahashi, et al., Global digital museum: multimedia information access and creation on the Internet, Proc. of ACM DL'98, pp. 244-253

- [3] Kounosu, K., et al., Outline of Network Library System in AFFRIC, Digital Libraries, No.10, pp.67-76, July 1997, http://www.DL.ulis.ac.jp/DLjournal/No_10/8-tzhaya/8-tzhaya.html (in Japanese)
- [4] Electronic Library Pilot Project, http://www.cii.ipa.go.jp/el/el/index_e.html
- [5] Matsumoto, H., High Speed Network for Digital Libraries, Proc. of ISDL'95, pp.65-72, August 1995
- [6] Aozora Bunko (BlueSky Collection): <http://www.aozora.gr.jp/>
- [7] Shinozuka, F. and Kuriyama, M., Digital libraries among national universities in Japan: an overview, The Journal of Information Science and Technology Association, Vol.49, No.6, pp.284-189, June 1999 (in Japanese)
- [8] Adachi, J., NACSIS Electronic Library System: Its Design and Implementation, Proc. of ISDL'95, pp.36-41, August 1995
- [9] Adachi, J., Dissemination of Japanese Academic Journals over the Internet, Proc. of ISDL'97, pp.32-35, November 1997
- [10] Oyama, K., et al., Construction of a Distributed Online Journal Editing System, Proc. Of ISDL'99, September 1999
- [11] Kando, N., Cross-Lingual Informaion Access and its Evaluation, Proc. of ISDL'99, September 1999
- [12] Imai, M., et al., Design of a Digital University Library: Mandala Library, Proc. of ISDL'95, pp.119-124, August 1995
- [13] Atarashi, R.S., et. al., Building Digital Library System -NAIST Challenge-, Proc. of ISDL'97, pp.28-31, November 1997
- [14] Atarashi, R.S., et al., Distributed digital lecture archive system: model and implementation, Proc. of ISDL'99, September 1999
- [15] Nagao, M., Multimedia Digital Library: ARIADNE, Proc. of ISDL'95, pp.1-4, August 1995
- [16] Kurohashi, S. and Nagao, M., Digital Library System at Kyoto University, Proc. of ISDL'97, pp.170-175, November 1997
- [17] Nakano, S., The Construction Plan of the Kansaikan of the National Diet Library and Digital Library, Proc. of ISDL'95, pp.81-86, August 1995
- [18] Taya, H., Strategic Plan for the Century 21 of the National Diet Library: Toward

- Establishing the National Electronic Heritage, Proc. of ISDL'97, pp.163-169, November 1997
- [19] (Reports on digital libraries)
<http://www.ndl.go.jp/projects/dento/doc/index.html>
- [20] Sakamoto, H., International Library of Children's Literature and Its Digital Library, Proc. of ISDL'99, Sept. 1999
- [21] Hiraoka, H., et al., Digital Library System at University of Library and Information Science, Journal of Information Processing and Management, Vol.42, No.6, pp.471-479, September 1999 (in Japanese)
- [22] Tabata, K., Prefice, Digital Libraries, No.1, pp.1-2, August 1994 (in Japanese)
- [23] Sakaguchi, T., et al., A Browsing Tool for Multi-lingual Documents for Users without Multi-lingual Fonts, Proc. of ACM DL'96, pp.63-71, March 1996
- [24] Maeda, A. et al., A Multilingual HTML Document Browsing System for Clients without Multilingual Fonts, Trans. Of IPSJ, Vol. 39, No.3, pp.802-809, March 1998, (in Japanese)
- [25] Maeda, A. et al., Viewing Multilingual Documents on Your Local Web Browser, Commn of ACM, Vol.41, No.4, pp.64-65, April 1998
- [26] Sugimoto, S., et al., Experimental Studies on an Applet-Based Document Viewer for Multilingual WWW Documents – Functional Extension of and Lessons Learned from Multilingual HTML, Proc. of 2nd ECDL, pp.199-214, September 1998
- [27] Dartois, M., et al., Building a Multi-lingual Electronic Text Collection of Folk Tales as a Set of Encapsulated Document Objects: an Approach for Casual Users to Browse Multi-lingual Documents on the Fly, Proc. of 1st ECDL, pp.215-231, September 1997
- [28] Dartois, M. et al., Building a Multi-lingual Electronic Text Collection of Folk Tales as a Set of Encapsulated Document Objects: an Approach for Casual Users to Browse Multi-lingual Documents on the Fly, D-lib Magazine, October 1997,
<http://www.dlib.org/dlib/october97/sugimoto/10sugimoto.html>
- [29] Baker, T., Dublin Core in Multiple Languages: Esperanto, Interlingua, or Pidgin?, Proc. of ISDL'97, pp.8-35, November 1997
- [30] Baker, T., Languages for Dublin Core, D-Lib Magazine December 1998,
<http://www.dlib.org/dlib/december98/12baker.html>
- Note:** Proceedings of ISDL'95 and ISDL'97 are accessible at <http://www.DL.ulis.ac.jp/{ISDL95,ISDL97}/proceedings/>.
- URLs of organizations referred**
1. National Museum of Ethnology: URL
<http://www.minpaku.ac.jp/>
 2. National Institute of Japanese Literature: URL
<http://www.nijl.ac.jp/>
 3. National Center for Science Information Systems (NACSIS): <http://www.nacsis.ac.jp/>
 4. University of Tokyo Library:
<http://www.lib.u-tokyo.ac.jp/>
 5. Hokkaido University Library:
http://www.lib.hokudai.ac.jp/index_e.html
 6. Kyoto University Library:
<http://www.kulib.kyoto-u.ac.jp/>
 7. Tsukuba University Library:
<http://www.tulips.tsukuba.ac.jp/>
 8. Tokyo Institute of Technology Library:
<http://www.libra.titech.ac.jp/>
 9. ULIS Library: <http://www.ulis.ac.jp/library/>
 10. NAIST Library:
<http://dlw3.aist-nara.ac.jp/index-j.html>
 11. Kobe University Library:
<http://www.lib.kobe-u.ac.jp/>
 12. NTCIR:
<http://www.rd.nacsis.ac.jp/~ntcadm/index-en.html>
 13. National Diet Library:
<http://www.ndl.go.jp/>
 14. JIPDEC (Digital Library Project):
<http://www.dlib.jipdec.or.jp/>
 15. IPA: [http://www/ipa.go.jp/](http://www.ipa.go.jp/)
 16. Agriculture, Forestry and Fisheries Research Council secretariat (AFFRC),
<http://www.affrc.go.jp/>