National Diet Library Newsletter

Reorganization of the NDL in FY2002

This is an abridged translation of the article of the same title in the NDL Monthly Bulletin No. 494

At the same time as the establishment of the Kansai-kan of the National Diet Library and the full opening of the International Library of Children's Literature (ILCL), the National Diet Library was reorganized on April 1, 2002.
Preceding this reorganization, the Research and Legislative Reference Bureau had been reorganized on April 1, 2001, in order to deal with the rise of demand from the Diet to improve legislative research services. For details, please see the *National Diet Library Monthly Bulletin,* No. 482, May 2001 (text in Japanese).

As a result of this reorganization, the NDL now consists of the Main Library (Tokyo Main Library with one bureau, five departments and the Detached Library in the Diet, and the Kansai-kan with one division and two departments) and two branch libraries (the ILCL and the Toyo Bunko) and 27 Branch Libraries located in the executive and judicial agencies of the government. For the new Organization Chart of the National Diet Library, please see *here*.

1. Tokyo Main Library

From materials to functions

Previous, some departments had been organized according to their collections (Books, Serials, and Special Materials Department). This time the Library was reorganized according to its function (acquisition, cataloging, storage, and public services) to improve its services for the Diet, executive and judicial agencies, and the general public.

（1）A Bibliography Department has been established to integrate the cataloging services that had formerly provided by several departments.

(2) A Public Services Department has been established to manage basic materials, books and serials, and provide comprehensive public services.

（3）A Reference and Special Collections Department has been established to provide appropriate reference service that meets today's needs,

Integration and enhancement of planning and coordinating function

Administrative aspects of library cooperation such as planning, coordination and negotiation that had been undertaken by the Library Cooperation Department, were integrated into the Administrative Department to enhance library cooperation functions. The Planning and Cooperation Division is in charge of these affairs. The Branch Libraries

Division that supports the Branch Libraries placed in the executive and judicial agencies of the government was also transferred to the Administrative Department.

Coping with the Information Technology Society

To undertake the construction of the Information Technology Society that is an important national policy, the Electronic Library Development Office, Planning Division, Administrative Department, was reorganized as the Digital Information Planning Office, Planning and Cooperation Division, Administrative Department. The Office coordinates digital information service projects that are being undertaken by several divisions such as the Kansai-kan and cooperates with both domestic and international advanced institutes in this field.

Enhancement of acquisition function

To deal with the wider demand for library materials and the diversification of information media, the cataloging function was separated from the Acquisitions Department. The acquisition and management function of library materials was effectively reorganized and the new Acquisitions Department established. It consists of the Acquisition Administration Division, the Domestic Materials Acquisition Division, the Foreign Materials Acquisition Division and the Preservation Division.

Integration and enhancement of bibliographic service

The newly established Bibliography Department integrates bibliographic service and provides it to the Tokyo Main Library, the Kansai-kan, and the ILCL. It consists of the Bibliographic Control Division, the Domestic Monographs Cataloging Division, the Foreign Monographs and Non-book Materials Cataloging Division, and the Serials Cataloging Division. In the Serials Cataloging Division, the Japanese Periodicals Index is also produced.

Comprehensive promotion of public services

The newly established Public Services Department coordinates public services operated in the Tokyo Main Library, the Kansai-kan and the ILCL and aims at providing total and effective public services. It consists of the Reader Service Planning Division, the Photoduplication Service Division, the Book Service Division, the Periodicals Service Division, and the Electronic Resources Division. This last provides not only audiovisual materials but also electronic journals and offline electronic materials in the library.

Enhancement of reference service

The special materials rooms of the Reference and Special Collections Department were reorganized to clarify their themes and aim at providing not only the Library's collections but also a wide range of information on those subjects. The Reference and Special Collections Department consists of the Reference Service Planning Division, the Business, Science and Technology Division, the Humanities Division (including library science and maps collection), the Modern Japanese Political Documents Division, the Rare Books and Old Materials Division, and the Newspapers Division.

In the Research and Legislative Reference Bureau, the Parliamentary Documents and Official Publications Division was newly established and provides statutes, parliamentary documents, government documents, official publications and reference books on law and

politics to the Diet members and the general public.

2. Kansai-kan

Establishment of the Kansai-kan

The newly established Kansai-kan of the National Diet Library consists of the Administrative Division, the Collections Department and the Projects Department.

(1) Administrative Division: In charge of management, personnel, accounts and the facilities of the Kansai-kan.

(2) Collections Department: In charge of acquisition, cataloging and public services. It consists of the Public Services Division (including remote services), Asian Resources Division, and Acquisitions and Cataloging Division

(3) Projects Department: In charge of library cooperation and the digital library project in the Kansai-kan. It consists of the Library Support Division (producing the union catalog, providing services for the visually and physically handicapped, conducting library science research and training for librarians) and the Digital Library Division (digitization and services via the Internet, and acquisition of online electronic materials). The Digital Library Division also conducts research and development of the digital library.

3. The International Library of Children's Literature

ILCL website

Establishment of the Children's Services Division

With the full opening of the ILCL in May 2002, full services for children were also started. The Children's Services Division takes charge of this function. The ILCL consists of the Planning and Cooperation Division, Resources and Information Division and Children's Services Division.

National Diet Library Newsletter

National Diet Library's FY2002 Budget

This is an abridged translation of the article of the same title in the NDL Monthly Bulletin No. 493

The fiscal 2002 budget bill passed the National Diet on March 27, 2002. The total amount of the initial budget appropriated for the NDL is 26,267,919,000 yen for FY2002 (April 2002 – March 2003). A decrease of 4.5 billion yen under the FY2001 level is mainly because of the completion of the refreshment of the International Library of Children's Literature (ILCL) and completion of construction of the Kyoto Official Residence of the NDL.
In FY2002, after the opening of the Kansai-kan of the NDL and the full opening of the ILCL, the NDL will have enhanced library services with all three facilities. Therefore the budget has been allocated to various new measures.

The composition of the total budget is:

- Salary and personnel expenses: 39.2% (33.0% in FY2001)
- Acquisition of library materials: 8.1% (5.0%)
- Development of information systems: 14.5% (12.0%)
- Office expenses: 22.0% (10.8%)
- Building and facilities: 16.2% (39.2%)

The following are major expenditures of the FY2002 budget.

related articles: FY2001, FY1999, FY1997

1. Kansai-kan of the National Diet Library

Approximately 3.4 billion yen was budgeted for the final allocation to the building construction. The remaining construction such as exterior work will be finished in August and the Kansai-kan will be opened in October 2002. Other costs of the Kansai-kan excluding the cost of construction are below.

(1,000 yen)

| Cost of preparations (e.g. transfer of library materials from Tokyo to Kyoto) | |

and opening ceremony	690,000
Operation of the Kansai-kan (e.g. library services and library cooperation project)	530,000
Acquisition of library materials	475,000
Administrative and maintenance expense for the Kansai-kan and the Kyoto Official Residence	1,134,000
Land acquisition cost for Phase 2 construction (approved from FY2002)	334,000

2. Tokyo Main Library

As part of the library materials and services are being transferred to the Kansai-kan, there will be a relocation of library materials, reading rooms and offices in the Tokyo Main Library. A three-year plan of renovation of the Catalog Hall and reading rooms will be undertaken in order to provide a new service that makes it possible to request materials and photocopies via computer terminal. As expenses for the renovation and reorganization of the Tokyo Main Library, approximately 270,000,000 yen was approved for FY2002. The total budget of the three-year plan will be approximately 2.5 billion yen.

3. International Library of Children's Literature

The ILCL, partly opened in May 2000, was fully opened in May 2002. After the full opening, the ILCL will upgrade user services for both children and researchers, start to provide services to school libraries, and increase electronic information of the bibliographic data and image data of its holdings. For operation of the ILCL, approximately 810,000,000 yen was budgeted.

4. Services for the Diet

Expenses for revising the Total System for Research and Information Services, through which the Library provides services to the Diet members via the Internet, were budgeted in order to deal promptly with requests from the Diet members.

Apart from the Full-text Database System for the Minutes of the Diet, expenses for construction of the Database System for the Statutes and Parliamentary Documents were newly authorized.

5. Electronic Library Service

Expenses for the Electronic Library Infrastructure System that enables the Tokyo Main Library, the Kansai-kan and the ILCL to provide services in an integrated way has been budgeted since the FY1998. For FY 2002, approximately 2,360,000,000 yen was allocated.

Expenses for the Electronic Library Services, which will be operated in the Kansai-kan,

have been included since the FY2000 supplementary budget and have been increasing content gradually. For FY2002, approximately 276,000,000 yen was allocated which will be used to operate projects such as the digitization of Japanese books printed in the Meiji era.

6. Acquisition of library materials

To provide information related to science and technology promptly, 70,000,000 yen for purchasing electronic journals provided in the Tokyo Main Library and the Kansai-kan was budgeted.

*(Accounts Division, Administrative Department)*

（1,000 yen）

Amounts of National Diet Library's FY2002 Expenditure Budget

(Item) National Diet Library	22,017,919
• Personnel costs	10,307,168
• Regular office expenses	384,567
• Library operations	485,544
• Office automation	3,797,212
• Acquisition of library materials	811,567
• Legislative research	314,337
• Editing and printing of basic bibliographies	31,665
• Acquisition of materials related to the U.S. Occupation of Japan	80,672
• Preservation of library materials	67,061
• Extra office expenses	164,223
• Library services for the blind	29,715
• Acquisition of materials related to Japanese Immigrants	10,797

• International Library of Children's Literature operation	809,853
• Annex of the Tokyo Main Library administration and maintenance	600,444
• Kansai-kan operation	2,828,672
• Construction of the Full-text Database System for the Minutes of the Diet	75,631
• Relocation of library materials, reading rooms and offices in the Tokyo Main Library	411,200
• Acquisition of science and technology materials	807,591
(Item) Expenses for facilities	4,250,000
• Expenses for the Kansai-kan facility	3,737,604
• Expenses for the renovation of the Tokyo Main Library	275,436
• Expenses for the Tokyo Main Library facility	236,960
Total	26,267,919

National Diet Library's Budget

National Diet Library Newsletter
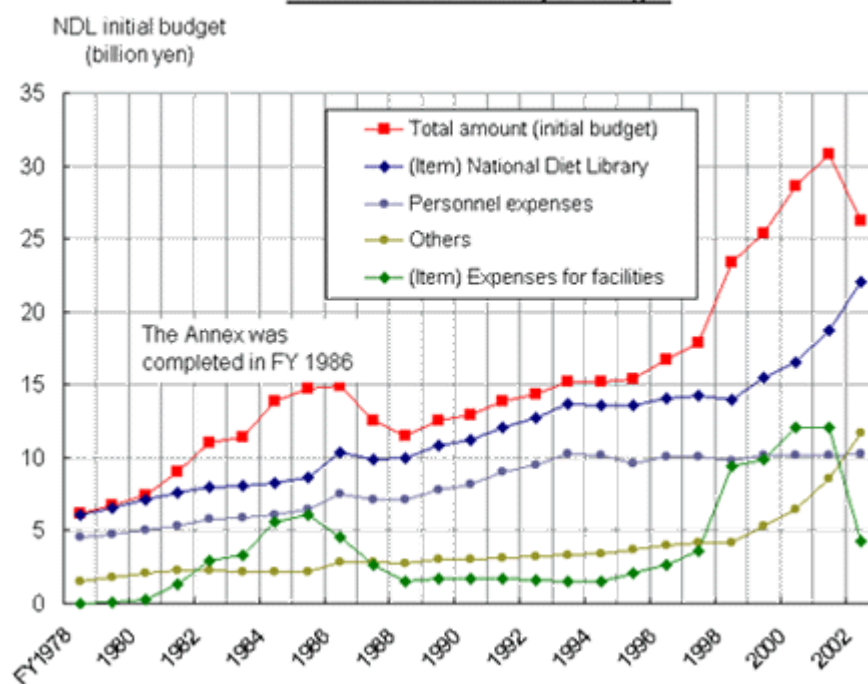
Full opening of the International Library of Children's Literature

The International Library of Children's Literature (ILCL), Branch of the National Diet Library, was fully opened on May 5, 2002. For the ILCL Website, please see *here* (text in Japanese). The ILCL was partially opened on May 5, 2000. As the first national library of its kind in Japan, it aims to acquire domestic and foreign children's literature and related materials, and make them available to the public. It is also promoting services appropriate to the 21st century, making use of digital library functions.

The Ribbon Cutting Ceremony

On May 4, 2002, the opening ceremony was held, inviting Her Imperial Highness Crown Princess Masako, Diet members, local elementary school students and other people. The commemorative exhibition "Friends in wonderland" is being held during May 5 thru September 14, 2002. On July 8, the commemorative symposium "From folktale to story" will be held, inviting experts from inside and outside Japan.

Building

The present building was constructed in 1906 and expanded in 1929. It is designated as a historic building by the Tokyo Metropolitan Government. While preserving the original form, it has been remodeled for its function as a children's library. The remodeling was completed and the library was fully opened in May 2002.

The design and structure of the interior and exterior of the valuable architectural heritage

has been preserved as much as possible. The building has also been made earthquake-resistant to protect it from major earthquakes.

Outline of the facility

Site	Ueno Park, Taito-ku, Tokyo
Structure	Brick building reinforced by steel frame; extensions are reinforced concrete
Size	One floor underground, three floors above ground
Ground area	Approximately 5,433 square meters
Total floor area	Approximately 6,671 square meters
Storage capacity	Approximately 400,000 volumes
Design and management	Kanto Regional Development Bureau, Ministry of Land, Infrastructure and Transport Tadao Ando Architect & Associates Co. Ltd. Nikken Sekkei Co. Ltd.
Preservation instructor	Katsuhiko Sakamoto, emeritus professor of Kobe Design University
Builder	Konoike Construction Co. Ltd.

Services

Anyone can use the ILCL, but in principle, only people over 18 years of age or older can use the Researchers・Resources on the second floor.

Resources

Resources and Information

In addition to the domestic books and documents acquired via legal deposit, the ILCL collects a wide range of literature and related materials from overseas. It also collects specialized information on children's literature and related materials.

Digital Library

Digital archive and digital museum functions support the ILCL, in addition to the PR and user guide functions of the ILCL website.

- Digital Archive: Specialized information is accumulated and provided in digital formats. e.g. Union catalog database of children's literature (in Japanese), Digital image database of children's literature
- Digital Museum: Virtual exhibition programs, which are based on picture books and other valuable images, are created and provided. e.g. Picture Book Gallery (in Japanese)

Reading On-site
Reading, photocopying and reference services are available in the ILCL.

Exhibitions and Events
Through exhibitions and events, the ILCL seeks to make children's books attractive and provides opportunities for children to be familiar with them.

Cooperation with School Libraries and Other Related Institutions
The ILCL lends selected sets of books and provides reference services to school libraries all over Japan. It also provides interlibrary loan, photocopy and reference services to public libraries and other institutions.

Collection (as of May 2002)

Children's books (including study books)

Books	Japanese	150,471 volumes
181,321 volumes	Overseas	30,850 volumes
Serials	Japanese	797 titles
817 titles	Overseas	20 titles

Research and reference materials on children's literature

Books	Japanese	9,296 volumes
10,496 volumes	Overseas	1,200 volumes
Serials	Japanese	660 titles
734 titles	Overseas	59 titles
	Newspapers	15 titles

Non-book materials

Still images, picture card shows	552 items
Cards	48 items
Music	209 volumes
Multimedia	236 items

Microfilms	36 reels
Microfiches	23,915 items
Optical disks	11 items

Materials for direct service to children

Books	Picture books	1,703 volumes
6,254 volumes	Literature	1,588 volumes
	Reference and information	2,963 volumes
Serials	Magazines	19 titles
23 titles	Newspapers	4 titles
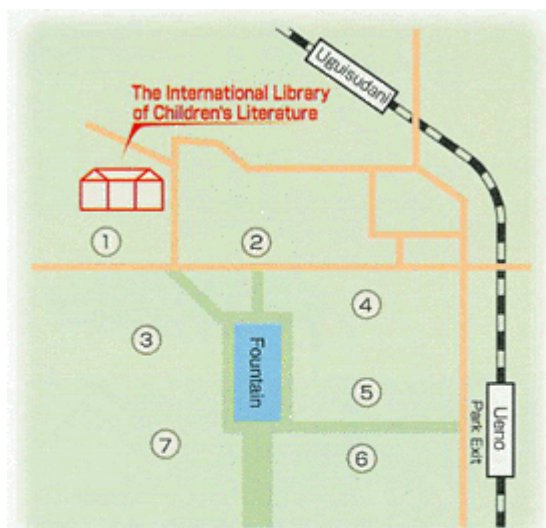
1. Tokyo National University of Fine Arts and Music
2. Tokyo National Museum
3. Tokyo Metropolitan Art Museum
4. National Science Museum
5. National Museum of Western Art
6. Tokyo Bunka Kaikan
7. Ueno Zoological Gardens

International Library of Children's Literature
12-49 Ueno Park, Taito-ku, Tokyo 110-0007
Japan
Tel: 81-3-3827-2053; Fax: 81-3-3827-2043
Web site: http://www.kodomo.go.jp
E-mail:info@kodomo.go.jp

up

National Diet Library Newsletter

Acquisition of school textbooks started

In April 2002, the National Diet Library began to acquire school textbooks issued in Japan. Because of limited storage, the presence of a textbook collection in the National Institute for Educational Policy Research and for other reasons, they have not been collected actively so far. However, the full opening of International Library of Children's Literature (ILCL) in May 2002 has solved the storage problem. Moreover, school textbooks have become more necessary for the ILCL to support the activities of libraries for children and provide services for researchers on children's literature.

With these movements in prospect, NDL consulted with the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Textbook Publishers Association of Japan and other related institutions about starting the acquisition of textbooks, putting the legal deposit system into practice. Consequently, with their consent, NDL launched the project in April 2002.

In FY 2002, the collection starts with elementary and junior high school textbooks and the corresponding guidelines for teachers. Textbooks for high school, schools for the physically challenged and mentally retarded children, except for Braille, and corresponding guidelines for teachers, will be added from FY 2003. At the moment, the whole textbook collection will be composed of publications that are supplied and used at schools in and after FY 2002.

The textbook collection will be available to the public in the ILCL after April 2003 because the new cataloging system will be in operation at the end of 2002. When the time comes, NDL will announce the details.

National Diet Library Newsletter

No. 125, June 2002
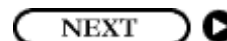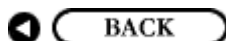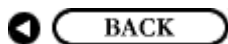
Report: Web Resources as Cultural Heritage: International Symposium on Web Archiving

This is a translation of the article of the same title in the NDL Monthly Bulletin No. 494

Table of Contents

For further information about the symposium (including the text of some of the papers), please see [here](#)

Introduction

Information resources on WWW continue to increase every day. How should we collect, archive, preserve them and provide permanent access? Based on these questions, the

International Symposium on Web Archiving was held on January 30, 2002, at the auditorium in the Annex of the National Diet Library, under the general theme of "Information on the World Wide Web as cultural heritage". There were 216 participants from inside and outside Japan. This paper is a brief report. The full proceedings of symposium are due to be published later.

Public Access to Digital Materials: Roles, Rights, and Responsibilities of Libraries by Dr. Brewster Kahle (Director, The Internet Archive)

The combination of a deep understanding and carrying of cultural heritage, as well as the best of the highest technologies, will make Japan a very strong player in building a digital library. In the information age when preservation and communication technology progress by leaps and bounds, some people think that libraries are no longer needed, that it is over. But this is surely wrong. Libraries have a significant role and responsibility in preserving our cultural heritage and providing universal access to digital materials. This may not be profitable, but it is still important.

The Internet Archive [1] has been collecting WWW in collaboration with other companies and groups. The Alexa Internet has been donating a copy of the WWW to the Internet Archive now for five years. The collection is over 100 terabytes (TB) in size, so it is a large collection. There are over 16 million different websites, with over 10 billion pages in this collection. The Alexa has also attempted to create metadata automatically, and then do subject indexing of websites.
As to offering material under copyright, after we talked informally (that is, off-the-record) with the head of the Copyright Office in the United States about one such collection, the following approaches were adopted. We would make the collection available at first, and if people asked us to take it down, we would take it down and explain about "robot exclusions" [2] [3]. So far, there's been no trouble. On rights issues, first we crawl websites and if the site owner contacts us we explain what it is that we are doing. The lesson for us was to try and make it available, and see what happens.

In addition, there is also the collection of television news. We made a collection from September 11 to 18, 2001, seven days, 24 hours a day, just one week, and took the news from 20 channels from around the world. It has been used to some extent by scholars and historians, and has been very positively received.
Today, I would like to present a gift to the National Diet Library from the Internet Archive, of a small sample of the Japanese websites that have been collected by the Internet Archive over the years. The collection is storing 20 million web pages from Japan, from 1996, 1997 and 1998. This is a collection that we hope will be useful in starting the Web Collection of the National Diet Library.

MINERVA: Mapping the Internet Electronic Resources Virtual Archive by Ms. Cassy Ammen (Reference Librarian, Library of Congress)

The Library of Congress's mission is to make its resources available and useful to the Congress and the American people and to sustain and preserve a universal collection of knowledge and creativity for future generations.

The MINERVA Web Preservation Project [4] was established to initiate a broad program to collect and preserve materials on the World Wide Web. The project team has developed a prototype system that provides access to an archive of web materials, and created representative collections of web sites related to the 2000 U. S. National Election and the September 11 Terrorist Attack.

We planned in our first pilot projects to take an open access approach, and select web materials both in a selective collection approach and in a bulk collection approach [5]. In the former, the library-staff select, collect, and catalog websites and then build a prototype access system to test and develop procedures for a production system. The latter was a project with the Internet Archive, and was a very productive collaboration. We are now in a new phase of a pilot, to work with a newly formed group known as Web-Archivist [6].

With regard to copyright, work is under way to interpret the Library's authority that it has already been granted by statutory authority, to move what we have done in our analog forms into the digital context. Interpretations should be consistent with the Library's established practices for non-digital materials, including our regular safeguards for rights holders' interests.

Next, the procedure for selection and collection is described. First, a small group of officers in the Library of Congress recommend the websites we should collect, based on various criteria. And then the website is downloaded using a mirroring program [2]. A snapshot is stored in an archive and additional snapshots are made at selected time intervals.

For our cataloging, we are using a system known as CORC (Cooperative Online Resource Catalog), which is part of the OCLC (Online Computer Library Corporation) cataloging system. MARC generated using the CORC interface would then be imported into the cataloging module of the Library of Congress's OPAC system, and then we would add subject headings, persistent identifier [7] and so on. For access, we developed a web based prototype system modeled on PANDORA Project in Australia [8]. You can search for websites by title, by subject, and by URL.

For preservation the Library of Congress has been appointed the lead agency among several government agencies to develop the National Digital Information Infrastructure and Preservation Program (NDIIPP). Another activity that the Library is beginning to participate in is a joint project through the OCLC and the Government Printing Office (GPO).

As for future explorations, we have to continue to study copyright, access issues, and legal deposit system, bulk collection, and establishment of selection criteria.

In the area of international cooperation, there are some interesting projects. For example,

the Electronic Resource Preservation and Access Network (ERANET) in Europe is a European Union endeavor, which is trying to preserve scientific information. UNESCO will be discussing, in its next budgetary year, preserving our digital heritage so that it might become a program in UNESCO [9]. Next fall in Rome, the European Conference on Digital Libraries (ECDL) [10] will be held.

If we are effectively to preserve for future generations the portion of this rapidly expanding corpus of information in digital form, we need to commit ourselves technically, legally, economically and organizationally to the full dimensions of the task. Failure to look for trusted means and methods of digital preservation will certainly exact a stiff, long-term cultural penalty.

up

Archiving the Web: The National Collection of Australian Online Publications by Ms. Margaret E. Phillips (Manager, Digital Archiving, National Library of Australia)

Recognizing that online publications are an intrinsic part of the documentary heritage, the National Library of Australia, together with a number of partners, is building the National Collection of Australian Online Publications.

Ensuring long-term access to online publications can be seen as a two-step process. First, the materials have to be identified, collected and made accessible in their current, or native, formats. That is the archiving process. Second, the materials have to be managed in such a way that they remain accessible as technology changes. This is the preservation process. The first step received most attention when we set up the PANDORA Archive in 1996. The preservation process has, however, been the growing focus of the Library's efforts as the archive has moved beyond the proof of concept stage to the operational National Collection.

The National Collection of Australian Online Publications is an extremely selective one, containing to date only 2,000 websites. Nevertheless, it already constitutes a representative sample of Australian web publishing by academic, government and commercial publishers, as well as community organizations. A number of the websites captured in the archive, including the official website for the Sydney Olympic Games, have already disappeared from the live Internet. Moreover, about one-third of the sites have been captured on multiple occasions [11]. The Collection now comprises almost 11 million files, and uses 320 gigabytes of storage. It is growing at about 500 new titles each year.

The National Library of Australia has deliberately pursued the selective approach to archiving for its advantages of quality control, and permission from publishers to archive and provide access. It has also developed a set of selection guidelines. However, in the future, we would like to supplement this selective collection with periodic snapshots of the entire Australian domain, working together with other agencies such as the Internet Archive.

The National Library of Australia and the National Film and Sound Archive, in a joint submission of the Copyright Law Review Committee, recommended amending the legal

deposit system to include non-print-based publications.

Besides, the National Library has been working with the Australian Publishers Association to develop a Code of Practice so as not to jeopardize publishers・commercial interests. In the year 2000, the Library engaged a consultant to provide advice on the direction it should take in relation to persistent identification of digital objects, and has implemented the guidelines. There are also a number of other authenticity issues [12], and problems related the archiving of databases [13].

In 2001, a very small trial migration [14] of 127 files was successfully undertaken. In the year 2000, RLG and OCLC invited the Library and a number of others to join an international working group charged with the task of proposing a draft international standard for preservation metadata. The working group has taken as its starting point the approaches to preservation metadata of the NEDLIB [15] or the Cedars [16]. The recommendations for the standard are expected in early 2002.

After five years we've learned our lesson. In the digital environment, it is advantageous to liaise and cooperate with a wider range of parties. Taking a practical approach and learning by doing has worked well for us, instead of just considering it theoretically. A team-based approach to devising and implementing policy and procedures enables us to draw on the expertise of a wide range of staff within the Library, and helps to motivate their commitment to the work. As a national deposit library, the National Library of Australia has clear responsibilities for collecting and preserving the documentary heritage of Australia in all its forms, and is ready to meet future challenges.

Danish Legal Deposit on the Internet by Ms. Birgit N. Henriksen (Head of Digitization and Web Department, The Royal Library, Denmark)

In 1997, the Danish legislation on legal deposit was modernized and updated, enabling the Royal Library to collect "static" (as opposed to "dynamic") work on Danish websites.

As for the process of legal deposit on the static work website, firstly, a publisher has to go to the legal deposit website and fill the metadata in a form. The metadata contains information about name, phone number, data format, whether any special program is used, and if user ID and password is required.

The library staff at the Danish Department determines whether the law covers the publication and if it does, downloads all files belonging to the work using our own harvester, verifies that all items are received and all hyper links are valid. Next, the work is catalogued and classified, and then finally it is transferred to the archival server. However we are not allowed to give access over the net to deposited digital works and this part of our system is therefore not public. The archived net publications can only be viewed at the reading rooms in the two legal deposit libraries.

In four years, we have collected 10,000 net publications, consisting of nearly 700,000 files and a total volume of only 23 gigabytes. Two-thirds of the publications are from the public sector such as the government or universities. The material collected mainly consists of

working papers, reports, scientific reports, guides, periodicals and newsletters.
For this selective web archiving, we do not need much hardware, so nearly all the costs are manpower. In order to decrease the amount of work, we skipped cataloging except for periodicals.

On the other hand, in bulk collections, we have to use techniques such as harvesting the entire Danish web space. It is necessary to harvest not only the ".dk" domain but also the publication placed in the ".com" domain. In addition, some material available to us as a user is simply not available for a harvester. This is, for instance, streaming contents [17] and materials with flash applications [18]. In order to archive dynamic publications, contracts or agreements had to be signed with the publishers. So we sent proposals for agreements to publishers, but it is to be regretted that only a few reacted to these proposals.

We wish to archive a broad range of the type of material that is to be collected as well as minimize the cost involved. Harvesting the entire sub-domain is suitable for this purpose and should be used to gather net material. But harvesting cannot solve all problems. It is still necessary to collect selectively for various purposes, and to use different archiving methods. We hope that in the future it will be possible to find a solution where materials can be freely accessible from the Internet.

Collecting and Archiving Web Resources and the National Diet Library by Ms. Machiko Nakai (Director, Electronic Library Development Office, National Diet Library)

At present, we plan to collect information resources on the network selectively, not by the legal deposit system. This is because the amount of digital resources is huge, and it is difficult to make publishers convert them into physical format and deposit to us.

As the standard of bibliographic description, we designed the National Diet Library Metadata Element Set in March 2001. This is based on the Dublin Core Metadata Element Set, and we adopted some original qualifiers that enable mapping to the JAPAN/MARC format.

Furthermore, we are developing the WARP (Web Archiving Program: provisional name), which is a total system of acquisition, archiving, maintenance, cataloguing, and provision of online resources. We designed a prototype for metadata entry up to March 2000, and are now developing a prototype of acquisition function.

The workflow is as follows: Selecting the resources to be acquired; Examining the structure of websites; Negotiating and contracting for acquisition with publishers; Specifying the unit of the information resources to be collected; Creating metadata; Setting harvesting and re-harvesting conditions; Trimming for removing the non-essential parts of the information; and Registering the individual object. The method of acquisition is assumed to be mainly a web-robot.

Web resources have the following characteristics: difficult to define unit of resources; easy

to be updated, changed and deleted; no hierarchical structure. The task is how we deal with these characteristics by the WARP. We are trying to set standards such as harvesting conditions and time interval of re-harvesting by making repeated experiments. As for Web-based databases, we plan to construct a navigation service based on metadata.

In early 2002, the Legal Deposit System Council will begin to discuss the legal deposit system for online electronic resources, including definition of the meaning of "publishing" and copyright issue. The deliberation will also be reflected in the WARP project.

up

Questions & Answers

After the lecture, questions and answers, moderated by Mr. Horoyuki Taya (Director, Foreign Materials Division, Acquisitions Department), were exchanged. The issues included reliability of collected information, long-term preservation, format for preservation, and security management.

An impressive comment on the reliability of collected information was made. That is, materials on the WWW are short-life and we have no time for considering their reliability, so the first thing to do is to collect and preserve them before they disappear.

up

Round-Table Discussion

The next day, a round-table discussion was held on systematic and technical problems at a working level. Most had the same opinion that both bulk collection and selective collection are necessary, and metadata should be assigned automatically. There was also an opinion that it is important to collaborate with the persons concerned such as publisher in its experimental stage.

up

Conclusion

This symposium on the ambitious theme, "Preservation of the Internet", was a success. It is probably the first time that this theme was publicly taken up in Japan. We hope that you can catch a glimpse of the future library of the 21st century.

up

Notes

1)The Internet Archive is a nonprofit organization that was founded with the purpose of collecting and preserving the information on the Web around the world. Its vast resources have been made available to the public through an interface called "The Wayback Machine" since October 2001. For details, see
http://www.archive.org

2) "Robots" are software programs that collect Websites automatically. Some are called "Web Robot", "Harvester", "Web-Crawler", or "Mirroring Program". Once you fix a URL as the starting point, the robot harvests a website as far as the depth designed, traversing hyperlinks recurrently. Collecting by robot is often called "harvesting".

3) "Robots exclusion" is a method that allows Web site administrators who do not want their site to be registered in search engines, to indicate their wish to exclude robots. The Internet Archive does not

collect the sites set up for robot exclusion, while the National Library of Finland, for example, includes such sites.

4) http://www.rlg.org/preserv/diginews/diginews5-2.html#feature1

5) The "bulk collection" approach is to harvest the information on the web by Web Robot in a wide range such as a whole country.

6) The Web Archivist is a research project to facilitate the archiving of specialized collections of web materials. For details, see http://www.webarchivist.org

7) "Persistent identifier" is a name assigned to information resource for assuring permanent access, which will remain the same regardless of where the resource is located. For example, the URN (Uniform Resource Name) is a kind of persistent identifier designed for improving the URL. The JP number and the ISSN are also authorized as its namespace.

8) http://pandora.nla.gov.au

9) http://unesdoc.unesco.org/images/0012/001255/125523e.pdf

10) The official name is the "European Conference on Research and Advanced Technology for Digital Libraries". In 2001, this conference was held in Darmstadt, Germany. Web archiving was one of the main subjects.

For details, see http://www.bnf.fr/pages/infopro/dli_ECDL2001.htm

About ECDL2002, see http://www.ecdl2002.org

11) Most of the information on the Web is revised frequently, so it is necessary to recollect regularly.

12) A document in electronic form is easier to tamper with than the conventional paper form. So it is a problem of how we can ensure reliability, authenticity, and admissibility as evidence.

13) Dynamic websites such as databases are called "deep web", which create and display the contents dynamically, and therefore cannot be harvested by web-robot. We also need to explore methods of collecting deep web.

14) "Migration" is the transfer of digital data from one system or format to another.

15) The official name is the "Networked European Deposit Library". The NEDLIB is a collaborative project of European national deposit libraries, and an attempt to design metadata for long-term preservation or as an original harvester tool.

For details, see http://www.kb.nl/coop/nedlib/

16) The official name is the "CURL Exemplars in Digital Archives". The Cedars is the digital preservation project promoted by the CURL (Consortium of University Research Libraries), UK, and other organizations.

For details, see http://www.leeds.ac.uk/cedars/index.htm

17) "Streaming" is a method by which a sound or movie file is sent from the server and executed at the same time.

18) "Flash" is a software that allows you to create web-contents including sound, graphics, and animation.

* Last access to referenced URLs above: May 10, 2002.

(Digital Library Division, Projects Department,
Kansai-kan of the National Diet Library)

National Diet Library Newsletter

Selections from NDL Collection
Seiro Bijin Awase

Title:	Seiro Bijin Awase
Designer:	Suzuki Harunobu (1725?–1770)
Date:	1770
Publisher:	Funaki Kasuke (et al.)
Description:	5 volumes, multi-color woodblock printed book, 27.0 cm × 18.2 cm
NDL call no.	WA32-5

The *Seiro Bijin Awase* is one of the finest multi-color woodblock printed books in Japan. It was published in 1770. The book depicts 166 courtesans of Yoshiwara, Edo (now part of Tokyo) and put the courtesans' names, the names of the brothels where they worked, and haiku in the background.

The *Seiro Bijin Awase* shows beautiful courtesans and their pastimes;

(picture1)
From the right: A courtesan called "Wakana" doing calligraphy exercises (page 1-cho-ura), a courtesan called "Ukifune" looking at a picture (page 2-cho-omote), vol.1

(picture 2)
Two courtesans called "Shizuhata" (right) and "Shizukaze" (left) playing cards (page

41-cho-ura , page 42-cho-omote), vol.3

(picture3)
Two courtesans called "Matsukaze" (right) and "Hatsuito" (left) holding a tea ceremony, (page 61-cho-ura, 62-cho-omote), vol.5

(picture 4)
Two courtesans called "Himegiku" (right) and "Katsuragi" (left) playing instruments (page 77-cho-ura, 78-cho-omote), vol.5

The *Seiro Bijin Awase* consists of five books in five volumes. The original title slip on which the book title and the volume title were printed is attached to in the center of the front cover (pic.5). Each volume is compiled according to the subject of the haiku: volume 1 for sakura (cherry blossoms), volume 2 for hototogisu (little cuckoo), volume 3 for tsuki (the moon), volume 4 for momiji (the beautifully tinted autumn leaves) and volume 5 for yuki (snow).

At the end of volume 5 is the imprint, which shows the publication date, the publisher's name and the etcher's name (pic.6). Though the designer's name is not directly mentioned, Suzuki Harunobu is thought to have designed the illustrations of the *Seiro Bijin Awase* as his name is mentioned in the preface. Harunobu was a leading ukiyo–e artist in the Meiwa period (1764–1771, part of

the Edo Period) and he was very good at depicting young beauties. He is also said to be the pioneer of nishiki-e (multi-colored prints), having produced a number of works of printed art.

(picture 5) Front cover of vol.4

(picture 6) Imprint at the back of vol.5

The *Seiro Bijin Awase* is a priceless masterpiece of Japanese art. It also possesses historic value in the development of printing techniques in Japan. In addition, the courtesans' behavior, the clothes they wear and the objects they use are worth observing because they show us many aspects of the Japanese culture. The NDL designated the *Seiro Bijin Awase* as a "Rare book". You can view all the pages of the *Seiro Bijin Awase* at the Rare books of the National Diet Library and the Rare Books Image Database (Japanese only).

National Diet Library Newsletter

◀ (BACK)
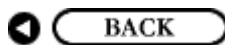
National Diet Library Monthly Bulletin
（Kokuritsu Kokkai Toshokan Geppo）

No. 495, June 2002

- Full opening of the International Library of Children's Literature (ILCL)

    - Summary of the commemorative events
    - Addresses given at the ribbon cutting ceremony
    - Addresses given at the reception
    - Exhibition commemorating the full opening of the ILCL: Friends in wonderland -from folktale to story-
    - May 4 for us: Comments from children
    - Message from the Director-General of the ILCL

- Library and information science in the 21st century and the future of the library: A meeting of researchers
- Materials recently designated as rare and semi-rare books: Report of the 33rd committee on the designation of rare books
- Partial amendments to the National Diet Library Law
- Processing "intangible resources": Trends in processing online electronic publications in Europe *by Ikuko Shiraishi*
- Notes on foreign parliamentary and governmental documents: looking into the G7 countries at the NDL. Part 9: France (3)
- Lecture "Japan in American children's literature: From Commodore Perry to today"

No. 494, May 2002

- Reorganization of the NDL in FY2002
- Report: Web Resources as Cultural Heritage: International Symposium on Web Archiving
- 5th Preservation and Conservation Training Program
- Notes on foreign parliamentary and governmental documents: looking into the G7 countries at the NDL. Part 8: France (2)

◀ (BACK)