Dynamic Transformation of the National Diet Library – From the Researcher's Library to the Library for All –

Kenji UETSUKI

Director

Digital Information Planning Office

Planning Division, Administrative Department

National Diet Library

Japan

Meeting:	**117 National Libraries**
Simultaneous Interpretation:	**Yes**

Abstract

Today it is undeniable that digital library services have become a reliable means of providing information. National libraries, among others, must make the maximum use of this new means, because the traditional means of providing services are limited by the physical restrictions which arise from national libraries' mission of preserving their nation's cultural heritage. The new ways of providing services have shown us how it is possible for us to navigate our patrons to whatever information resources they need by breaking the constraints that have long existed. On the other side, digital information resources have won the title for the most promising medium carrying information on politics, economics and culture. National libraries, which have been a nation's base to preserve physical artifacts such as paper-form materials, now must take digital information resources into consideration. We must surmount high hurdles in order to develop legal systems that support the

preservation of digital information. The National Diet Library (NDL) has been working hard to tackle this task, aiming to preserve information resources on the Internet. To pursue these aims, we have been working on the following three objectives:

1) To construct a major digital archive of the nation as an expansion of the library's remote service

2) To enhance a navigation system which guides users seamlessly to various collections and information resources wherever they are regardless of kinds of media; to develop a one-stop portal which guides users to collections and resources

3) To preserve information resources on the Internet

1. Introduction

<Restrictions on access to the original materials>

The NDL, as the only legal deposit library of Japan, has a mission to collect the nation's heritage and pass it on to future generations. For this mission to be fulfilled, the NDL has to put restrictions on the use of materials. It usually happens that only one copy of a publication is deposited with the NDL, and the collection is gathered and stored in one place. Also within the limited human and physical resources, service hours are limited to a certain period of time, and therefore the library usually closes at night and on holidays. With these restrictions, users are naturally limited to certain groups, such as librarians and researchers. There was also a restriction by users' age. From the point of preservation, we had long regarded the NDL as the last resort.

This is our library's past, and the situation has not changed much in terms of restricted access to the original paper-form materials. In this regard I believe that the same may apply to other national libraries. However, there has been a radical change going on at the NDL since the opening of the Kansai-kan in October 2002, when we started full-scale digital library services.

2. First Transformation

<Availability of remote access>

By providing full-scale bibliographic information on the Internet, it has become possible for remote users to check to see if a certain item is held by the NDL and if the item is available for loan or photocopy. It has also become possible to request loan or photocopying of the materials via the Internet. Full-scale digital library services began when the NDL started to provide digital contents on the Internet, which can be browsed directly by remote users without having to visit the library.

Since 2002 when the NDL started providing the NDL-OPAC (National Diet Library Online

Public Access Catalog)[1], the NDL and its collections have become more visible and transparent to remote users. That is, it has become easy for remote users to obtain information that had been long unavailable to them without visiting the library in person, such as information included in the catalogs located only in the library and about the availability of each material. The NDL-OPAC covers most of the library's collections, including monographs, periodicals, newspapers, doctoral dissertations, and other forms of materials, and it contains 14.4 million records. Half of the records contained, some 7 million records, are from the Japanese Periodicals Index, which indexes some 10,000 titles of Japanese periodicals, including both academic journals and popular magazines. All the records of the NDL-OPAC are available free of charge.

Registered users can make requests online for photocopies or interlibrary loans of the library's materials. In 2004 73% of the requests for photocopies were made online, and the registered users amounted to 64,000. User registration is free of charge. By introducing the user registration system, the number of photocopy requests more than tripled from 2000 to 2005. The main driving force for this increase is the rising number of individual users who use this remote service. This is the first step in the transformation of our library's service in which the main pillar has shifted from on-site service to remote service.

Another step in the transformation is the beginning of the so-called "digital library service" through which users can directly browse library materials from remote places. In the first project, the "Digital Library from the Meiji Era,[2]" we digitized monographs published in Japan during the 45 years of the Meiji era (1868-1912) and held by the NDL, and have provided them in the form of digital images on the Internet since 2002.

Of the whole collection of the Meiji-era monographs held by the NDL, about 75% of the copyright were cleared, amounting to approximately 130,000 volumes. The project was completed in spring this year. A little less than 30% of the digitized works were confirmed as copyright-free. However, the remaining 70% of the works were found to be still copyrighted, but the copyright owners could not be located even though we made considerable effort. For these works of whose copyright owners we could not find, we obtained a compulsory license from the Commissioner for Cultural Affairs. This is a system based on the Copyright Law: a compulsory license is given without copyright owners' permission by depositing estimated copyright fee with the government so that we can provide the copyrighted works to the public. In April this year, the volumes digitized were doubled, and access to them was tripled.

---

[1] http://opac.ndl.go.jp/index_e.html
[2] http://kindai.ndl.go.jp/

To use rare and old materials, users need to submit an application in advance, and their accessibility is not high. Now, however, we have digitized them and made them open to the public in the Rare Books Image Database[3]. Now approximately 37,000 images of 861 titles are available.

In addition, we have made available digital exhibitions[4] on nine subjects so far. One of the exhibitions, the Birth of the Constitution of Japan, has been highly evaluated and even used as a school educational tool. The number of accesses to it amounts to 250,000 per month.

As a parliamentary library, it is our unique service to electronically provide full-text of minutes of the Diet[5]. What gives prominence to this database is that it includes not only the minutes of Plenary Sittings but also those of the Committees.

We have also digitized our publications that used to be published in paper form and made them available online. For example, the NDL Annual Report, public relations magazines such as the NDL Monthly Bulletin and the NDL Newsletter, and reports made for reference of the Diet deliberations, are available on the NDL website.

<Shift to borderless service>
Remote service using the library's collections and information on them mentioned above has become nationwide and has obviously narrowed the gaps between the regions. Internet service itself does not have national borders, though language diversity exists. Resources on the NDL website are used as a tool to study Japan by academic institutions world-wide. If users have the right equipment, they can access it wherever they are, living in a foreign country or traveling abroad on business. There is no border between countries to restrict access. With the remote service, users can also make use of our collections without time constraints. Accessibility has been clearly improved. The access counter of our website records over 15,000 per day, which is ten times as many as the number of on-site users. Of course, we do not need to set an age limit to use the website.

The behavior of users who visit the library in person has also been changed by the influence of remote service. For example, the number of requests for materials that cannot be provided on-site is reduced, while the number of materials one person uses in a day increases. That means the remote service is not only useful for remote users but also has a good effect on allowing visiting

---

[3] http://rarebook.ndl.go.jp/pre/servlet/pre_com_menu.jsp
[4] http://www.ndl.go.jp/en/gallery/index.html and
http://www.kodomo.go.jp/gallery/digi/index_e.html
[5] http://kokkai.ndl.go.jp/

users to use the library effectively.

It can be summarized that the remote service has had a dramatic effect in improving accessibility of our services. This is the first stage of the digital library service without walls that we have been working hard for the last decade to establish. We believe that we have successfully reached it.

3. Further possibilities

Here I will talk about our plan for the next three years.

Two years and a half ago, we drew up a basic policy, which highly evaluates the functions of the digital library service. The policy, called the NDL Vision 2004[6], describes our mission, role, and the priority areas. Creation of digital archives and improvement of access to information resources are clearly advocated in the priority areas. Under the Vision, the National Diet Library Digital Library Medium Term Plan for 2004[7] sets forth the direction in which the digital library service will be oriented over the next five years, and presents the framework required for realizing such service. The direction of the construction of digital archives we have been undertaking follows the Plan. As the means of circulation of information change, the role of library itself also changes. To respond to the demands of the time for a change of national library, we set three main goals as follows:

1) To construct a major digital archive of the nation as an expansion of the library's remote service
2) To enhance a navigation system which guides users seamlessly to various collections and information resources regardless of whether we hold them or not and regardless of kinds of media; to develop a one-stop portal which guides users to collections and resources in an integrated manner
3) To preserve information resources on the Internet

This is a declaration that the NDL, which formerly served as the last resort for researchers, should serve as a national library for all people in Japan and should function as a part of the world library. We are trying to break out of our old shell to transform ourselves dynamically.

---

[6] http://www.ndl.go.jp/en/aboutus/vision_2004.html
[7] http://www.ndl.go.jp/en/aboutus/elib_plan2004.html

3.1 Creation of digital library

<Digitization>

We will continue to intensively digitize our holdings published in Japan and make them open to the public. The digitized collections are a major part of the NDL Digital Archive together with the archived Internet information resources I will explain in more detail later. We finished the digitization of books published in the Meiji era and are preparing to digitize the next part, 90,000 books published in the Taisho era (1912-1926). The digitization of rare and old materials that do not require copyright clearance will also be expanded.

In addition to the provision of information on paper-form materials such as periodicals index and bibliographic data, and the function of receiving requests for copying service via the Internet, we will steadily provide our collections electronically, the easier parts first.

In addition, we have developed the Collaborative Reference Database Project[8] to help users have access to expert knowledge on information. This database has been available on the Internet since 2005, providing accumulated data such as reference examples, manuals to find resources on specific subjects, special collections held in each library, and profiles of member libraries, with the collaboration of 390 libraries including public, university and special libraries. This information is to support reference services in libraries and the research and study of general users. As of April 2006, a total of 21,172 records are registered, of which 10,131 are open to the public.

3.2 Construction of a digital archive portal

Accumulation of digital contents is important for the library of course, but it is not sufficient enough. The value of a national library will be judged by its potential to provide integrated information to the individual user wherever they exist. We have guided users to our databases with GUI (Graphical User Interface) so far, specific to each database. Our navigation service has been prone to be limited mainly to resources of related institutions such as libraries and academic organizations. But now we are planning to integrate and provide various database of the library first, and then expand it to integrate information of many digital archives outside, in the way we can achieve it automatically. For its realization, we are preparing to construct a portal site. In this site users who have searched a specific material will be guided to digital materials in one-stop access or, for instance, to the photocopy service for paper-form materials. Users now expect seamless access both to paper-form and digital-form materials, and what they need is not limited to information inside libraries. The coverage of search should be expanded beyond the library walls to archives, commercial databases, book stores, secondhand booksellers, museums and so

---

[8] http://www.ndl.go.jp/jp/library/collabo-ref.html

on. In this way we are aiming at constructing a mechanism which serves as a bridge between users and the information needed, at one stop as far as possible. We call this service, mentioned as the second main goal, a digital archive portal. We are now in the stage of constructing and providing a prototype system[9] before starting full-scale construction. This pilot system has been developed since fiscal 2004 in order to examine required specifications for the full-scale operation. For example, by entering a search word you can find a list of results from multiple databases such as the Digital Library from the Meiji Era, NDL-OPAC, National Archives of Japan and Japanese Periodicals Index and then jump to those primary and secondary sources.

<Standardization of interface>

What is the mechanism to realize these functions? In general, most of the websites and informational services used to have their specific service interfaces through browsers. There are not so many of them which have an application program interface enabling a cross search for other service providers for their integrated information services. If each institution provides application program interface under the same specifications, and if it automatically provides metadata of each content in a form processable by computers, various service providers and portals can make use of it to develop value-added services in response to the variety of needs and interests of users.

We are aiming to realize a digital archive portal by promoting and generalizing the system like this, because it is obvious that there is a limitation in the way of preparing linkage software according to each specification. Promoting common interface to the data providers, we expect that it will be possible for us to have a wider linkage with other institutions such as libraries, academic organizations, archives, book stores, secondhand booksellers, and museums, and also we expect that users will be happy to have one-stop access to the contents.

Every institution navigating information broadly will need a mechanism like this sooner or later. The important point is to collect metadata automatically from each database. Regardless of whether the targeted content is in the surface or in the deep web, if only the existence and the location of that content are indicated mechanically, that can be used for to achieve a one-stop portal. This mechanism has been already realized for secondary sources, through the Z39.50 protocol in library environments. In addition to this, the OAI-PMH(Open Archive Initiative – Protocol for Metadata Harvesting) protocol is used these days for institutional repositories of academic organizations including universities, and Web services such as SRW(Search/Retrieve Web Service) seem to be developing widely in the future. On the practical side, more and more

---

[9] http://www.dap.ndl.go.jp/home/

sites are putting information on the web in a simple way using RSS(RDF Site Summary or Really Simple Syndication). The important key for the future is the standardization and dissemination of the mechanism mentioned above that will realize an integrated search of NDL's materials and a navigation system for users to the broader information they need.

\<Bridge to the preservation of digital information\>
In addition, I have to point out that the digital archive portal has a great deal to do with the future preservation of information resources on the Internet. In principle, content accessible automatically in a portal should be able to be collected by web robots even if it is in the deep web. Therefore, promotion of the one-stop portal useful for daily information seeking will also contribute directly to a system for preservation. In this point, we think this project is also an important steppingstone toward the archiving of resources on the Internet in the future. It may be better that the information we are now using will be conserved for the next generation. As for network preservation, I will talk about it in the next section.

Until now, the NDL has regarded itself as the last resort. This will not change, but we are also aiming to be from the beginning a library for all, which will be used everywhere.

3.3 The preservation of information resources on the Internet
Lastly, I should mention the common issue national libraries in the world already share. It is the preservation of digital information that is in danger of disappearing, such as information resources on the Internet. No one denies any more that digital information has been realized as the most promising medium carrying information on politics, economics and culture. National libraries, which have been a nation's base to preserve physical artifacts such as paper-form materials, now must take digital information resources into consideration. Although we must surmount high hurdles in order to develop legal systems that support the preservation of digital information, it is our important mission as a national library to preserve information created electronically and guarantee its accessibility.

The NDL has been working on a project in order to preserve websites since 2000. It is called WARP (Web Archiving Project)[10], collecting and providing information by the permission of each website. It contains electronic journals and websites of government agencies, some local agencies prior to their mergers, public corporations, institutions, universities and some events which have probably been deleted after their closing, such as exhibitions. The oldest collection is that of 2000.

---

[10] http://warp.ndl.go.jp/

We found, through a simple investigation conducted on last July, that, among the periodicals deposited with the NDL, over 250 titles of periodicals have been converted into digital form from paper form. While this is an example of what used to be published in the paper media and is now distributed in digital media, the forms of distribution have assuredly and significantly been changed in the information-communication society. It has become an undeniable fact that information on the Internet and the manner of its utilization is, apart from its advantages and disadvantages, becoming a large foundation for businesses and culture. However, digital information has characteristics that can be easily updated, deleted and moved. These characteristics pose major challenge for preserving digital information. We think the NDL should reflect these characteristics and take a role as a mirror to secure the survival of this generation's knowledge into the future.

As is well known, UNESCO adopted the "Charter on the Preservation of the Digital Heritage"[11] in 2003 and stressed its importance to governments. In response, the NDL has set the following as main targets for future projects: preserving and accumulating of information resources on the Internet, web archiving and providing it.

<Comprehensive acquisition>
In the WARP project, having dealt with many websites, some problems have been pointed out: complicated relations of rights that arise from an enormous amount of copyright clearance; what digital information should be preserved for future generations; and how to collect digital information which cannot be collected automatically. Accordingly, our library has been continuing preparations for the full-scale acquisition of information on the Internet trying to prepare for a legal system. We have three ways. The first is bulk harvesting, which collects the surface web contents comprehensively by using crawler, following links to links by restricting their copyrights. Although collecting all the information on the Internet in Japan might be worth doing, it is said that the Internet carries much illegal information as well. Therefore we assume that we will practically limit the scope of the application of the bulk harvesting to the websites of highly public organizations. In addition, to avoid repressing publishers' motivation to freely express and publish their ideas as a result of the NDL's acquisition non-consensual with publishers, we will adopt the policy of accepting publishers' claim not to collect their publications. Publishers can also make a claim to delete what we have collected of their publications and not to make them public on the Internet. At the same time we will stop providing some kinds of information in cases where we find a fact or have claims from the victims of child pornography,

[11]

obscenities, or other encroachments of human rights.

\<Selective acquisition\>

The second way is to collect websites selectively. By the above-mentioned way, we are unable to collect useful information existing in the purely private domain. Therefore, we are going to collect selectively in regard to the websites of corporations, private organizations and individuals. It also enables us to consider the frequency of acquisition separately.

The third way is to collect individual works that deserve to be held on the electric bookshelves from information resources hidden in the deep web and undetectable automatically by crawler. It includes methods to attach the metadata, which makes the individual works searchable, and enables us to render them with the future equipments. With this methodology we hope that we can play a role in the international cooperation to preserve digital information, with a view to preserving digital journals as has been done by the National Library of the Netherlands[12]. The National Library of the Netherlands and the NDL made an arrangement for comprehensive cooperation in August 2005, and we have made efforts to cope with the common problems through international cooperation.

We are still engaged on the development of legal systems at present and that is why we should mainly collect digital information selectively, though we have kept up our efforts for the development of technology, consensus building and acquisition aimed at preserving the comprehensive digital information resources of Japan. There are many problems and tasks in constructing preservation systems and developing technologies for rendering digital information, which means especially that mutual cooperation among national libraries is essential.

4. Conclusion

In order to change and improve so as to be the library for all, as the communicator of information seamlessly throughout the eras, while fulfilling our mission, we have to work on the following:

1) To preserve and accumulate resources of digital information in response to the changing information environment
2) To ensure that people can share and enjoy the fruits of new information and communication technology by utilizing it

For the former, I overleap explanations because it had been already mentioned in many papers before. For the latter, standardization of the interface and promotion of its dissemination is

---

[12] http://www.kb.nl/bst/beleid/bp/2002/opslag-en.html

necessary. This requires a cooperative relationship with relevant institutions. Services need to be developed, taking into account the various kinds of libraries and regional libraries, for the expansion and equality of the opportunities to share new services.