National Diet Library Newsletter

No.192, December 2013

NDL's Digital Collection and Service for Information Access

Naoko Harai

Library Counsellor

Acquisitions and Bibliography Department

This article is extracted from the lecture presented

at the OCLC Asia Pacific Regional Council 5th Membership Conference in Bangkok, Thailand

on Monday-Tuesday, October 7-8, 2013.

First, I will introduce some of NDL's Objectives that are related to today's theme.

Next, I will introduce, mainly focused on internet services, construction of the digital collection, thirdly, provision service of digital collection, and at last, provision of bibliographic data.

Contents

1. Introduction

NDL has published on its homepage, "Mission and Objectives 2012 – 2016". It was formulated and released in July 2012 for five years up to 2016. Its English version is also available on the homepage.

"Our Mission"

"Our Mission" shows the basic mission to be accomplished by the NDL.

"The National Diet Library extensively collects and preserves publications and other materials/information from inside and outside the country as a knowledge base and cultural infrastructure; assists the Diet and provides library services to the executive and judicial branches and the public; supports the creativity of the nation and contributes to the development of democracy."

"Objectives 2012 – 2016"

"Objectives 2012- 2016" consists of six objectives. Of all the objectives, the second, the third and the fifth are related to today's theme.

Objective 2: Acquisition and Preservation

"We will further improve the legal deposit system and strive to build comprehensive collections of domestic publications; besides print publications, we intend to acquire and preserve a variety of materials/information, both digital and non-digital, as cultural assets."

Digital materials are also treated as to be acquired.

Objective 3: Information Access

"We will prepare universally user-friendly environments and measures, through digitization of materials, improved search methods and others, in response to new information environments, so that our holdings are easily available and sought-after information can be found promptly and precisely."

Under the Objective there are some strategic goals. These are "Digitize library materials / provide Internet access", "Access to information resources in Japan and abroad", and "Improve utility of bibliographic information".

Objective 5: Archives of the Great East Japan Earthquake

"We will continue to develop the Great East Japan Earthquake Archives in order to convey to future generations the records of and the lessons learned from the unprecedented disaster."

As you know, the Great East Japan Earthquake on March 11th 2011 was a very powerful experience for us. As a national library, the NDL must collect and preserve resources about the earthquake for the future.

2. Construction of Digital Collection

NDL's Construction of the digital collection has two pillars. One is the digitization of the NDL's collection, holdings. And other one is the collection of online information resources.

Digitization of the NDL's collection

We began the digitization with the rare books and old materials. The result was provided through the "Rare Books Image Database" on the Internet in March 2000. Next the NDL focused on

digitizing the books published after the Meiji Restoration of 1868. The outcome has been available in the "Digital Library from the Meiji Era" on the Internet since October 2012.

However, not all has been smooth sailing; the digitization had to be done on a meager budget. So the annual increase of the digitized contents was by only about ten thousand. The digitized contents held by the NDL reached only around 150,000 items by 2008.

A major turning point came in 2009. The supplementary budget for the purpose of stimulating the economy allocated 12.7 billion yen to the NDL for digitization.

In addition, in 2009, the Copyright Act was amended. It has brought a change to NDL's digitization. Since then the NDL, the sole deposit library in Japan, can digitize its holdings for the purpose of long-term preservation and provide the digital images within the NDL's premises, without permission of the rights holders. Before the amendment, the Copyright Act allowed libraries to digitize only severely deteriorated items, so the NDL had always carried out digitization with permission.

Current state of digitization of the NDL's collection

Both by the amendment of the Copyright Act and the enormous budget in 2009, the NDL could make great progress on the road of digitization. As a result, by now, 90 thousand items of Rare Books and Old Materials, 895 thousand Japanese books, one million 50 thousand Japanese periodicals and 140 thousand doctoral dissertations are already digitized. In total, including other materials, two million 250 thousand materials have been digitized.

Web archiving

Next I will talk about the other pillar of NDL's digital collection, web archiving.

NDL launched the Web Archiving Project, WARP, to collect websites in November 2002. The project was initially piloted experimentally based on permission of copyright holders such as website owners. With regard to the public sector's websites, the NDL worked on law amendment to collect and preserve them without permission; and the NDL Act, amended in 2009 and put into force in April 2010, enables the NDL to do so.

Transition of the accumulated amount of the harvested data

The number of collected websites is 7,466 items and that of harvesting is 55,749 times in total. The data amount was 231TB as of the end of March 2013.

Collection of digital books and periodicals

Unlike public sector websites, collection of information resources released on the Internet by private publishers, companies, individuals, private universities and research institutes still requires permission from those owners.

Among the rapid spread of digital books and periodicals, the collection, preservation and provision of those issued by the private sector needed another law amendment. The NDL pushed for the legal amendment, and thus the NDL Act was again amended in June 2012 and the amendment act came into force in July 2013. This amendment obliges publishers of digital books and periodicals to deposit their publications with the NDL. Since the NDL was still in negotiation with publishers and rights holders for the collection, preservation and provision of paid digital books and periodicals, the amendment act exempts paid publications and DRM (Digital Rights Management System)-protected publications for some time, and the target for legal deposit is any of the following among free and non-DRM publications published (released) on the Internet and other transmission media, by the private sector.

In addition, the Academic Degree Regulations (Ordinance of the Ministry of Education, Culture, Sports, Science and Technology) were amended in April 2013 so that a doctoral dissertation must, as a rule, be published through the Internet. The NDL had, so far, relied on donation, not deposit, from universities etc. for collection of doctoral dissertations in paper form, throughout Japan. From FY 2013, it collects and preserves doctoral dissertations as digital contents in the same manner as digital books and periodicals.

3. Access to Digital Collection

As a third point, I talk about our service on provision of digital collection. And I explain about NDL Search as an integrated search service adding navigation service to other organization's holdings.

Provision of digitized collection

For Japanese books, Japanese periodicals and doctoral dissertations of all digitized materials, NDL needed copyright clearance such as checking the year of author's death to confirm expiration of copyright or discovering copyright owners to obtain necessary permission. As a result, 344 thousand items of Japanese books, 5 thousand items of Japanese periodicals and 15 thousand items of doctoral dissertations are provided on the Internet. Including 70 thousand items of Rare

National Diet Library Newsletter

Books and Old Materials, 20 thousand items of official gazettes and so on, 470 thousand items in total are available on the Internet as of June 2013.

Archiving Websites

Archived websites are also available on the Internet.

NDL Digitized Contents

Here you can look at NDL digitized materials. You can select categories from the bottom lines.

WARP

You can see archived websites on the WARP page.

Publications collected from the Internet (Internet Resources)

You can see NDL Digitized Contents through selecting 'Internet Resources' born-digital publications such as white papers, annual reports, PR magazines collected from the websites of national agencies, local government agencies, independent administrative entities, and universities.

NDL Search

"NDL Search" is a new search service. This service aims to be an access point from which you can use plentiful information from public libraries, archives, museums, institutes of academic research and the NDL.

The search target contains 205 databases, amounting to some 73 million data. NDL Search has functions to search using as a key word contents information and, if it is an NDL publication, all words within the material. You can use the NDL Search for forthcoming books and newly acquired materials. The NDL Search also gives access to various databases from the search results, not merely linking to databases of holding institutions but also showing how to access the information or materials you need as much as possible.

You can access NDL's digitized materials and archived websites using the NDL Search.

Function of NDL Search

NDL Search has a multilingual translation function. The interface is available in English, Chinese and Korean. You can search for materials written in other languages (English, Chinese and Korean) by the function that translates search keywords and results.

NDL Search has additional functions to make use of search results: provision of permanent URL of bibliographic data, RSS feeding of search results, downloading results data for RefWorks and so on.

NDL Search provides RDF (Resource Description Framework) metadata obtainable through API.

4. Provision of Bibliographic data

Finally, I will explain about our bibliographic data provision as NDL's Service for Information Access.

Traditional service

NDL provides bibliographic data through various routes using different formats. An old provision service, printed JNB, "Japanese National Bibliography" ceased publication in 2007, and its HTML version had been provided from 2002 until 2011.
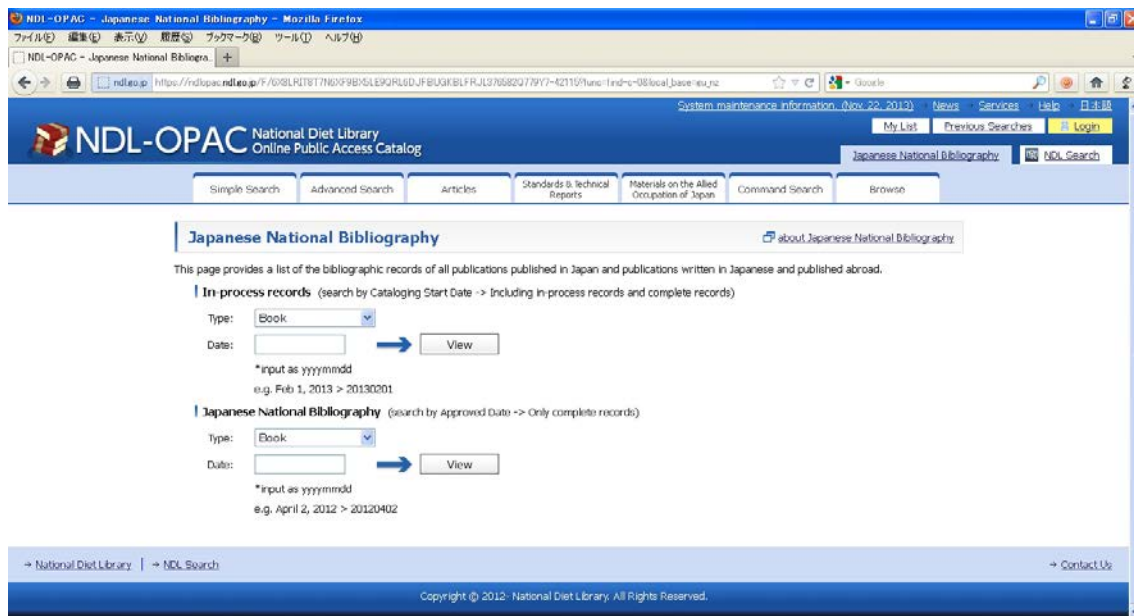
The machine readable national bibliography, JAPAN/MARC, has been distributed since 1981, and continues now. There is also CD-ROM version added with search function, J-BISC or JMBISC. They are distributed through distributors, not directly from NDL.

On the Internet, we provide NDL-OPAC that is an online public catalogue of NDL's holdings. NDL-OPAC includes bibliographic data within JNB and out of JNB when NDL holds their material. As of March 2013, NDL-OPAC includes 11 million data of books, serials, newspapers etc. Eleven million data of the Japanese Periodicals Index are available through NDL-OPAC. Registered users can also use the copying service.

New service: NDL-OPAC

NDL has developed some new services of bibliographic data provision in the recent years.

NDL-OPAC has changed its bibliographic format from JAPAN/MARC to MARC21 in January 2012. At the same time, a sub-menu, "JNB provision service" started. By this menu, you can get data sets within the JNB range by downloading. As NDL-OPAC is our online public catalogue, of course you can get bibliographic data out of the JNB range.

National Diet Library Newsletter

<< Japanese National Bibliography>>

You can also get newly added JNB data in process.

Download format can be selected from MARC format and text format.

New service: NDL Search

NDL Search is an integrated search service, and here you can get bibliographic data. Its format is DC-NDL, NDL's original metadata based on the Dublin Core.

NDL started to collect e-books and e-magazines in July 2013 and will start providing part of eBooks and eMagazines included in JNB from next April through NDL Search. For these data provision, RSS feeding of NDL Search is used.

NDL Search provides API so that external systems can utilize its functions.

New service: Web NDL Authorities

Web NDL Authorities is a new service of NDL. This service provides all authority data of NDL, which includes NDL's author authority data and NDLSH, NDL Subject Headings, on the Internet. In 2010, we started to provide NDLSH data as Web NDLSH and in 2012 enlarged it to name authority data and changed its name to Web NDL Authorities.

Web NDL Authorities: Number of records

Personal Name	790,176
Family Name	2,180
Corporate Name	184,955
Uniform Title	3,989
Geographic Name	27,476
NDLSH Topical Term (Incl. Topical + Subdivision)	101,919
NDLSH Subdivision	323
Total	1,111,018

<<Web NDL Authorities: Number of records as of March 2013>>

Here is the number of the authority records provided by Web NDL Authorities: Personal Name: 790 thousand, Family Name: 2 thousand, Corporate Name: 180 thousand, Uniform Title: 4 thousand, Geographic Name: 27 thousand and NDLSH Topical Term: 102 thousand. Totally 1 million 100 thousand records are available.

Web NDL Authorities: Features

Web NDL Authorities is applied to semantic web. It adopts RDF for description, gives URI for each of the headings, uses the SPARQL for search. You can download all NDLSH data.

Web NDL Authorities data are updated daily automatically. This service has RSS feed function for newly added NDLSH data.

You can also follow from Web NDL Authorities to VIAF, LCSH or Wikipedia through links.

International Provision of data: OCLC

NDL started to provide bibliographic data to OCLC in 2010. At first JAPAN/MARC (M), Monographs data of JNB was provided four times a year, and from this April, the provision has been monthly. Now JAPAN/MARC (S), Serials data of JNB is also available at OCLC Webcat. Next we will provide Japanese Periodicals Index data.

Number of records are 4 million 500 thousand of JAPAN/MARC (M), 150 thousand of JAPAN/MARC (S), and 10 million 900 thousand of the Japanese Periodicals Index.

Now all are updated every month.

International Provision of data: VIAF

NDL has participated in VIAF, Virtual International Authority File, and now provides the total of 1 million records. Update is monthly. As I already explained, Web NDL Authorities data are linked to VIAF data.

NDL has an active policy to promote the use of our bibliographic data. Please try them.