National Diet Library Newsletter

No. 249, February 2023

Contents

国立国会図書館
National Diet Library, Japan

Selections from NDL collections

"*Ushibori*" by KAWASE Hasui From *Kawase Hasui Hangashu*

Published in 1930. 25.7 × 38.5 cm, one sheet
Owned by the National Diet Library, Japan
*Available in the NDL Digital Collections.

KAWASE Hasui (1883–1957) was an artist who specialized in drawing Japanese landscapes during both the Taisho (1912–1926) and Showa (1926–1989) eras. Travelling extensively across Japan, he would sketch local landscapes, from which he later produced a number of woodblock prints. His travel-oriented style of realism earned him the moniker "travelling poet."

Ushibori depicts a winter landscape of a quiet, snowy night in a riverside village, where a small boat ferries a lone guest across the river. Straight ahead of the boat, a warm light from a window is splashed across the dark surface of the water. This single instance of a warm color creates a beautiful contrast with the solemn silence of the indigo that dominates the painting. This is an example of *aizuri*, a technique that had been used in *ukiyo-e* since the Edo period (1603–1868) and is characterized by a gradated indigo coloration set off with an occasional second color, usually red or yellow. It was mainly used for landscapes and pictures of beautiful women (*bijin-ga*). In *Ushibori*, the indigo of the snowy sky was created using a technique called *zarazuri*, in which the strokes of a *baren*, a rubbing pad used for printing woodblocks, create a contrast with the transparent surface of the water.

Most of the Hasui's works were published by WATANABE Shozaburo (1885–1962), a publisher who specialized in the use of a method for creating woodblock prints that was common during the Edo period, in which three specialists collaborate to produce a finished product. In this method *eshi* (artists who painted the original pictures), *horishi* (carvers who sculpted the woodblock), and *surishi* (printers who applied ink to the surface of the cut woodblock and printed them onto paper) work together to create an original woodblock print. Shozaburo was particularly careful to communicate the intentions of the *eshi* to the *horishi* and the *surishi*, which enhanced the overall artistry of woodblock prints. Shozaburo's insistence on these three specialists collaborating while making full use of traditional techniques held a particular import for Hasui's works.

The beautiful, indigo snowscape of *Ushibori* was created using techniques and skills that were originally developed for *ukiyo-e* of the Edo period.

*Ushibori is a place in the modern-day city of Itako in Ibaraki Prefecture.

(Translated by NOZAWA Asuka)

Articles by NDL staff

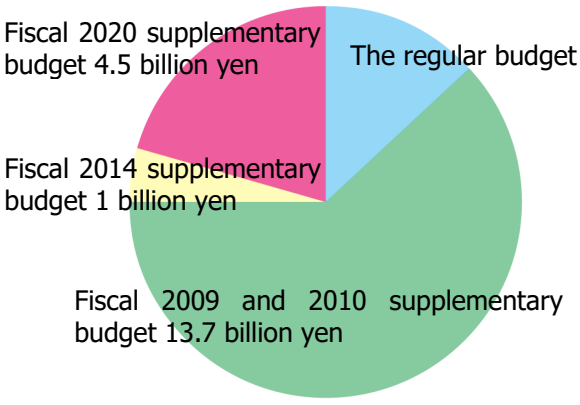
Digitizing Library Materials at the NDL (part 1)

Digitization Promotion Office, Digital Information Planning Division, Digital Information Department

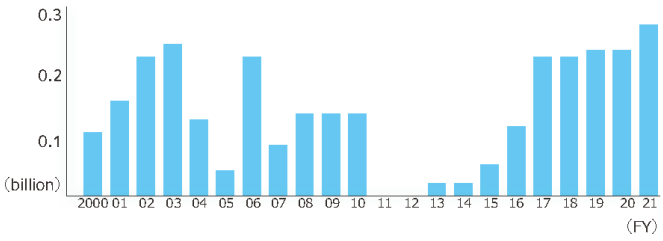
This article is a translation of the article in Japanese in NDL Monthly Bulletin No.733 (May 2022)

For the expansion of a digital archive of its holdings, the National Diet Library was allocated 6 billion yen in funding from the fiscal 2020 supplementary budget. Based on the budget, the NDL used 4.5 billion yen to outsource the digitization of some 300,000 books, roughly 45 million digital images, during fiscal 2021. It was the single biggest digitization project since the first mass digitization project in 2009 and 2010, which produced some 900,000 digitized books, based on 13.7 billion yen in funding from supplementary budgets.

The digitization process itself might seem to many people to be similar to making copies on a copy machine. But when the scale of the project is such that you must digitize 300,000 materials in a variety of forms, and that some of these are one-of-a-kind items that simply cannot be damaged, then the process becomes closely akin to mass production of industrial materials.
In this article, we present an introduction to little-known, behind-the-scenes operations used to digitize materials during the FY2021 project.



Past budgets for digitization



Allocations for digitization in the regular budget

01 Selection and preliminary arrangement of materials for digitization

Selection

The first step in the digitization process is the selection of the target materials.

We selected target materials based on the March 2021 *Basic Plan for Digitization of the National Diet Library's Collections 2021–2025*. We considered characteristics such as uniqueness, rarity, expansion opportunities for use, state of deterioration, and urgency of preservation as well as the demand for digitization both from the general public and academic society. Given that digitization would take place on a massive scale, we selected groups of materials, not individual materials, according to type and age.

Making a list
Calculating number of frames

Once the target materials were determined, they were checked for damage and added to a list of materials to be digitized. The final list contained 450,000 volumes in FY 2021, which is equivalent to a single floor of the stacks in the Tokyo Main Library.

Furthermore, the digitization contract is based on the number of frames, which generally comprise two facing pages on a single image. Thus, we have to calculate the approximate number of frames based on the number of pages and illustrations. Some old materials do not have enough bibliographical information, and some have special shapes, so calculations are done by trial and error. Also at this time we repair damaged materials.

02 Packing library materials

Materials on the list are taken from the stacks and packed for transport to the digitization vendor's facilities. To protect the materials from being soiled and damaged, they are placed into folding containers with cushioning. One folding container is large enough to hold about 25 individual materials. In FY2021, 18,000 boxes were packed. Since each box has a volume of 50 liters, the total volume was 900,000 liters or equivalent to a swimming pool about 25 x 36 x 1 meters in size.

Packing the materials in folding boxes

Empty bookshelves

03 Transporting library materials to the vendor

About 400 folding containers can be loaded onto a single, 10-ton truck. In FY2021, a total of nearly 100 10-ton trucks were used to transport materials.

Inventory control is performed to prevent the wrong boxes from being transported by mistake. Each box is marked with a barcode and checked one by one.

The containers are loaded onto a big truck.

04 Pre-digitization check

Digitization was outsourced to multiple vendors, each of whom has specialized skills and expertise. Each step in the digitization process is carried out on its own line.

Verifying the list

After the materials are unloaded at the vendor's facility, the first step is to confirm that the content of each box corresponds with the packing list. The list is checked every time materials are moved or digitized to ensure that nothing is lost during processing.

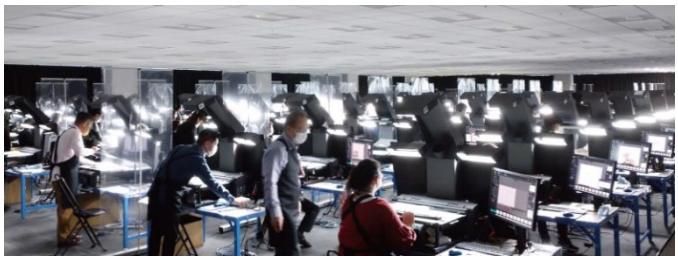
Checking the materials

The next step is to check the size and condition of the materials, decide which scanner to use, and to make a note of any points of attention during digitization.

To ensure that materials do not deteriorate while at the digitization vendor's facilities, they are stored in a controlled environment at the same temperature and humidity as at the NDL: 22°C and 55% RH. Also, the humidity at the digitization unit is controlled to always be within 10% of that of the storage unit. Temperature and humidity in both places are measured and recorded regularly.

Checking materials

05 Imaging



Digitization is performed using book scanners. 100 scanners all lined up in rows is a magnificent sight.

Aligning the pages under the glass plate with a ruler

Adjusting the height of the right and left plates for a thick book

Turning pages quickly yet carefully

Throughput of the digitization process is highly dependent on the condition of the materials, including its thickness, flexibility, and the presence or absence of fold-out diagrams. Normal production is about 500 to 1,000 frames a day per operator.

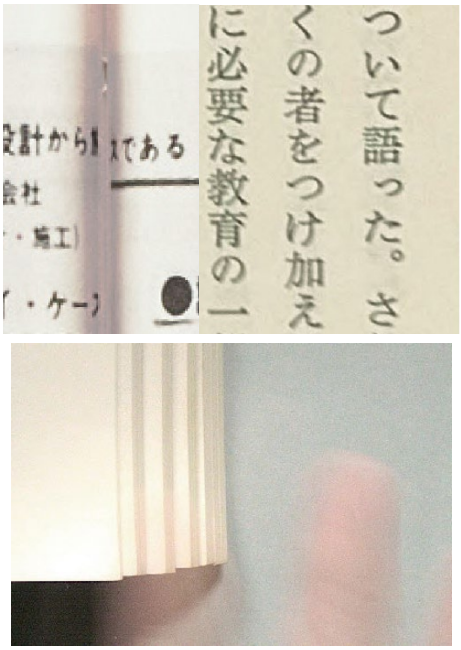
The digitization process itself is performed by hand and requires a great deal of experience. The material is held down flat with a glass plate to protect against damage as well as to ensure that even the characters near the binding are clearly legible. If the glass plate moves, the image will be blurred, yet productivity requires that the pages be turned at a good tempo.

06 Quality check and rescanning

The quality of images is checked, and any images that do not pass inspection are rescanned. This process often detects a significant number of errors that require rescanning.

Typical errors include skipped pages, duplicated pages, out-of-focus pages, or fingers visible in the scan.

The specifications call for an error rate of 0.5% or less. To achieve this, each vendor makes an effort to improve quality using their own expertise. Nevertheless, with some 45 million frames to scan, the occurrence of errors is unavoidable.



Typical errors: clockwise from top right, out of focus, fingers visible, and book not open enough

07 Making index data

Making digitized images available to the public also requires creation of index data for the table of contents as well as bibliographic data, including title, year of publication, and other metadata.

We used to type the table of contents, but in FY2021 changed the specification to give priority to digitization, and now we only record the name of files which include the table of contents. Owing to the enormous number of characters in the Japanese language, high-quality text input also requires a great deal of time and effort.

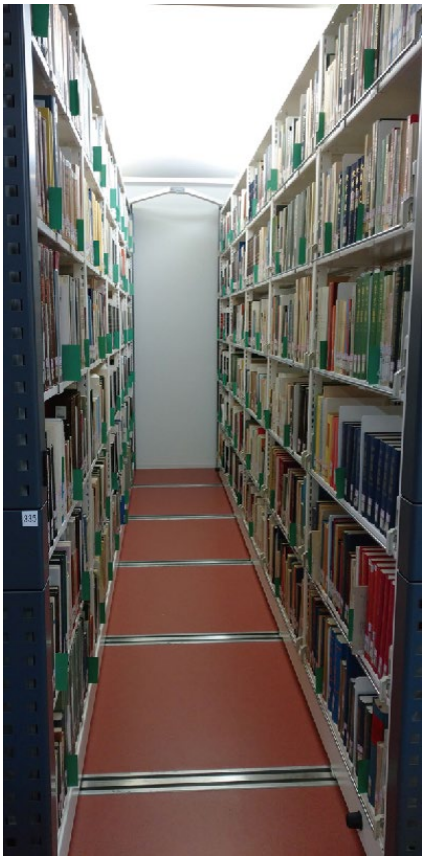
Looking at the scanned images, operators verify which frames contain the table of contents and input their frame numbers.

08 Project administration

Each vendor employs more than 100 people just to carry out the scanning, so thorough project administration is necessary.
Management systems are required to deal with large volumes of materials, data and metadata. Differences in scanning skill make it difficult to estimate productivity, so managing human resources is a major issue. In addition, during FY2021, the COVID-19 pandemic and the global semiconductor shortage made things hard going for this project.

09 Returning materials

After the digitization process is complete, the materials are returned to the NDL and put back on the shelves. Previously, we returned the materials to the stacks of the Tokyo Main Library, but since NDL policy calls for the originals of digitized materials to be transferred to the Kansai-kan approximately 400 km far from Tokyo, during FY2021 we transported the originals of digitized materials directly to the Kansai-kan.



At the Kansai-kan, the materials are unpacked, labeled with a location code, and stored in the stacks.

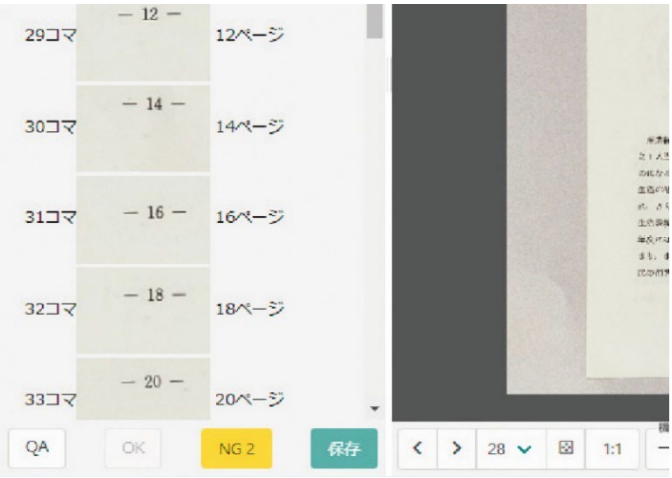
10 Delivery of digital images

The checked images are stored on hard disks, which are delivered to the NDL together with the metadata. Approximately 1.5 TB of data is generated for 10,000 materials.

Roughly 200 hard disks containing digitized images were delivered for this project during FY2021.

11 Inspection of digitized images

The digitized images are also inspected by the NDL.
In FY2021, we created a new system that enables us to control for the inspection operation by making lists of errors. In this system, the file format of all images and their correspondence to metadata are checked mechanically, with a visual inspection by humans performed for randomly sampled materials.



A screenshot of the inspection system. Only the page numbers of the images are checked to ensure there are no skipped pages.

12 Uploading and making digitized images available to public

The digitized images are made available to the public via the National Diet Library Digital Collections.
The digital images are delivered on hard disks to the NDL and gradually transferred to the NDL Digital Collections server at data centers. Therefore, even after the digitization is finished, it takes a bit of time before the digitized copies are available.
We are now developing a new system for the NDL Digital Collections to help resolve this problem.[1]

The storage devices in which digital images are saved

A new system for the NDL Digital Collections is now under development.

In-house digitization

As you can see, mass digitization is a kind of industrial production process, and outsourcing can only be reasonable for a large amount of homogeneous material. Smaller quantities of materials in varied forms are risky for both the client and the contractor. Therefore, if a material is difficult to photograph, deteriorated, or heavily used, the library staff may digitize it in-house. In FY2021, a media conversion room equipped with equipment for this purpose was established in the Tokyo Main Library.
The library has been photographing old children's books, government booklets, and special media packages such as laser disks.

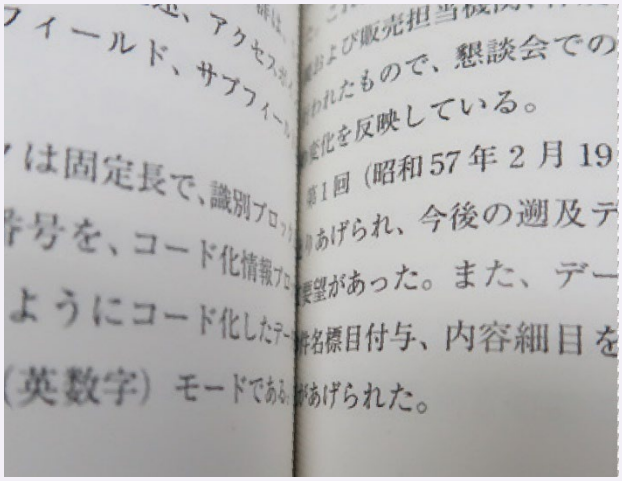
[1] The original article in Japanese was written using information available in May 2022. The NDL Digital Collections underwent a major renewal on December 21, 2022.

Ideas on digitization

Books vary widely in size, thickness, and binding. In order to digitize books as a clean and clear image, these differences need to be accommodated. Here are some examples.

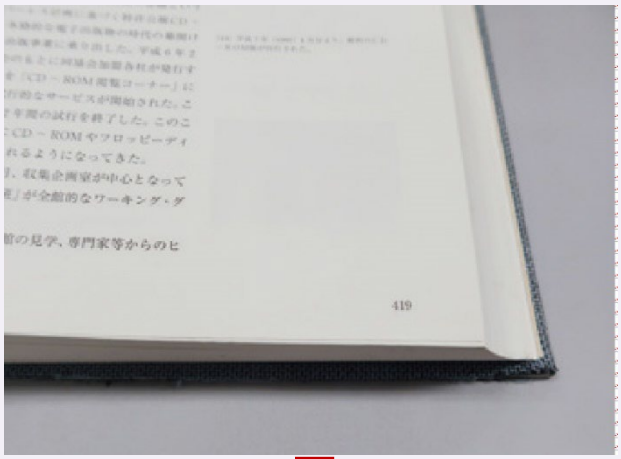
Material with binding defects

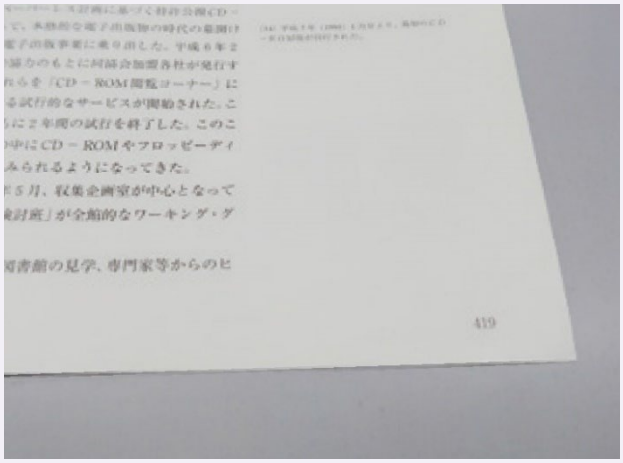
When text is hidden in the gutter, we loosen the binding before photographing. Since there are various kinds of methods of binding, we handle them in different ways depending on their condition.



Thick materials

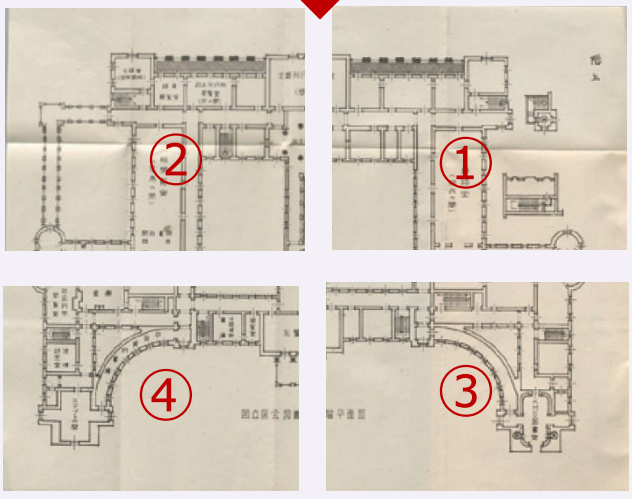
When thick materials are photographed as they are, their images often become hard to read. For example, they tend to be distorted and out of focus, or the size of margins near the cover and in the center of the document can differ greatly. To avoid these issues, we allow photographing by inserting a backing sheet for thick materials, to hide the fore edge and eliminate the three-dimensional effect.





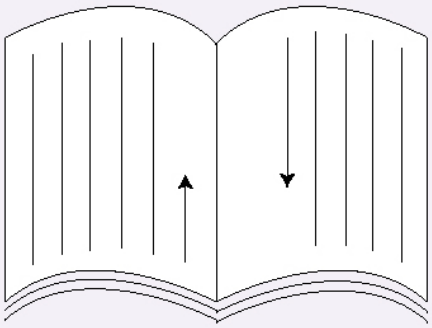
Folded pages

Sometimes, pages with illustrations, errata, etc. are folded in. We photograph them both folded and unfolded to show the condition of the original document. If the unfolded pages are too large to photograph at the same size as the main body, we divide the frame and photograph them with a little overlap. The order of the frames is based on the direction of page transitions in the main body.



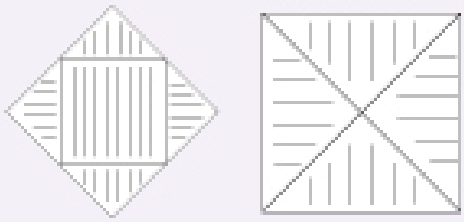
Upside down

Sometimes, the second half of a document is bound upside down as a design choice. For ease of reading, we align the direction of all the parts and photograph them. In addition, we add a note in the metadata remarks, "The second half of the document is bound upside down."



Materials folded into squares

We photograph everything on both the front and back sides, because both sides have information and they only make sense when looking at each stage of the folding.



Clipping as design

Sometimes there are cutouts on several pages, and the front and back pages can be seen peeking through when they are superimposed. We determine that the cutouts were made as part of the design, so we photograph each page individually as well as the overlapping pages in as many patterns as possible.

(Translated by HATTORI Mao and MACHIDA Sumika)

Some pictures courtesy of MUSASHI CO., LTD. and NIPPON EXPRESS.

Articles by NDL staff

The Journey to Digitization

Book Services Office, Book and Serials Division, Reader Services and Collections Department

This article is a translation of the article in Japanese in NDL Monthly Bulletin No.733 (May 2022)

When I explain to family and friends that I am in charge of digitization at work, they sometimes respond by asking if that means that my job involves placing books on a scanner to make data. But actually, my job is a little different from what they imagine. At the National Diet Library, since what we do involves the large-scale digitization of numerous books, most of the scanning process is outsourced. So, my section is involved in preparing the books for their journey to the digitization vendors, specifically, selecting and inventorying the target materials as well as repairing damaged ones.

One of the most important jobs we do is to survey the materials. We go into the stacks and check each book one by one to verify their content, size, and condition. Some of the books are big and heavy, some are more than a meter wide when laid open. There are times when several of us work together to lay the book open and check its condition.

After the results of the survey are collated and an inventory made, we use this information to decide which books should be digitized. In addition to the title, volume number, and other bibliographic information needed to identify the book, the results of the survey are also listed in the inventory. Many aspects of the digitization processes, including those not described in this article, are performed using the information found in the inventory list. An enormous number of books are digitized, and having an accurate inventory list is extremely important, since any errors in the list can result in books getting lost or otherwise disrupting the digitization process. It goes without saying that compiling the inventory list is a very painstaking process.

In addition to making inventory lists, another essential part of what we do is preparing the books for shipping. Books that have supplementary items have to be labeled to identify the content of the supplement, and if a book is damaged, we make simple repairs intended to enhance the appearance of the digitized image.

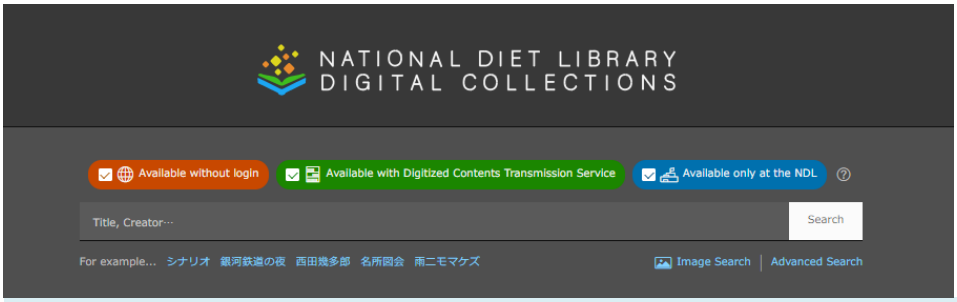
Although each of these tasks is relatively insignificant on its own, performing them all as part of a larger process has made me keenly aware of the wide range of materials that the National Diet Library holds. I remember once I was so captivated by a book with beautiful cloth samples pasted into it that I lost track of time. Another time, I found an intricately folded illustration of an anatomical model, which made me wonder how it would be scanned. Once, we even had a book with a ceramic plate embedded in the cover, and we had to cover the ceramic plate to protect it and the surrounding area from damage.

It's rather humbling to realize that the books I worked on with my own hands will become digital images for posterity, but it also makes me happy to know that books that have been read so many times that they are tattered and in fragile condition will become available for reading on computers, smartphones, and similar devices. Today, I am diligently packing books to ensure that they go in as good condition as possible on their journey to digitization.

(Translated by HATTORI Mao)

News from NDL

Renewal of the National Diet Library Digital Collections



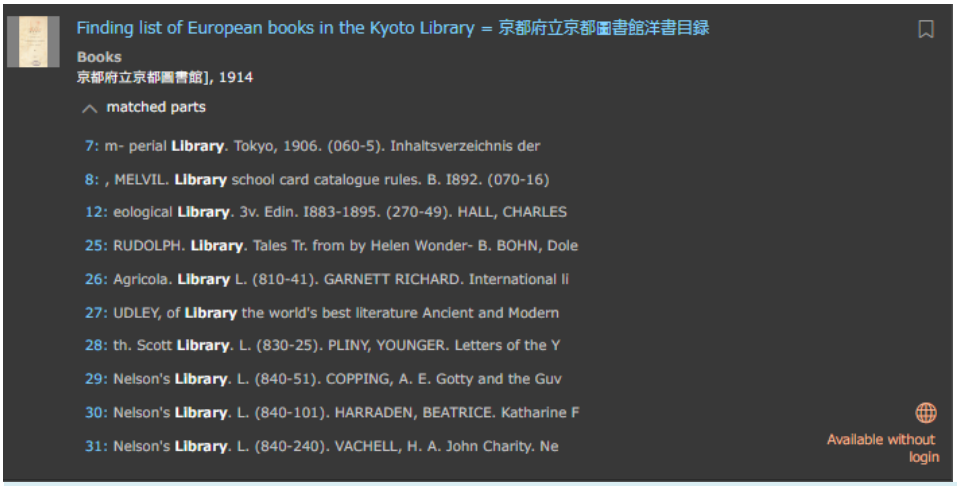
The top page of the renewed NDL Digital Collections

On December 21, 2022, the National Diet Library (NDL) renewed the National Diet Library Digital Collections (NDL Digital Collections), in which 3.11 million digitized materials and 1.5 million digitally acquired books and periodicals are available. Its search function and browsing page have become more convenient and useful. Please use the new NDL Digital Collections for your research, hobbies and entertainment!

Main contents of the renewal

1. More full-text-searchable materials

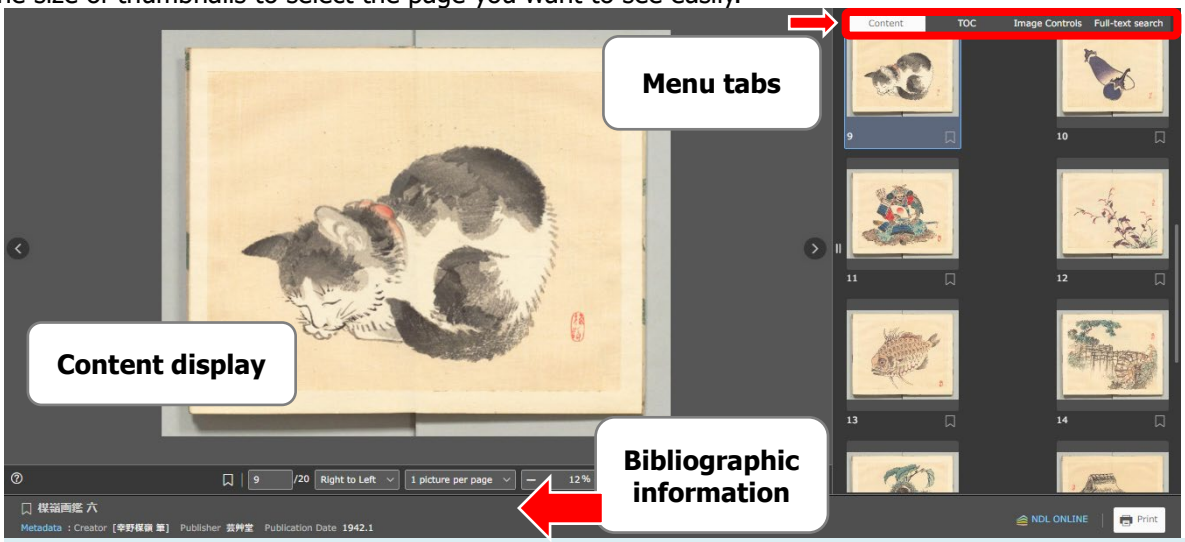
The NDL carried out text conversion for the books and periodicals which were uploaded to the NDL Digital Collections up until December 2020. As a result, the number of full-text-searchable materials has increased from 50 thousand to about 2.47 million. The results of full-text searches are displayed as a list, and you can move to the relevant point in the material directly.



An image of the display of full-text search results in the NDL Digital Collections

2. Improvement of the browsing page

Bibliographic information is displayed at the bottom of the page and the content display can be expanded. You can also change the size of thumbnails to select the page you want to see easily.



The browsing page of the NDL Digital Collections and menu tabs

3. Image search function

Image search function is newly available for searching the pictures from the books and old materials in the NDL Digital Collections whose copyright protection terms have expired. Using pictures from the NDL Digital Collections or other websites, or your own pictures, you can search for similar figures, illustrations and photos.

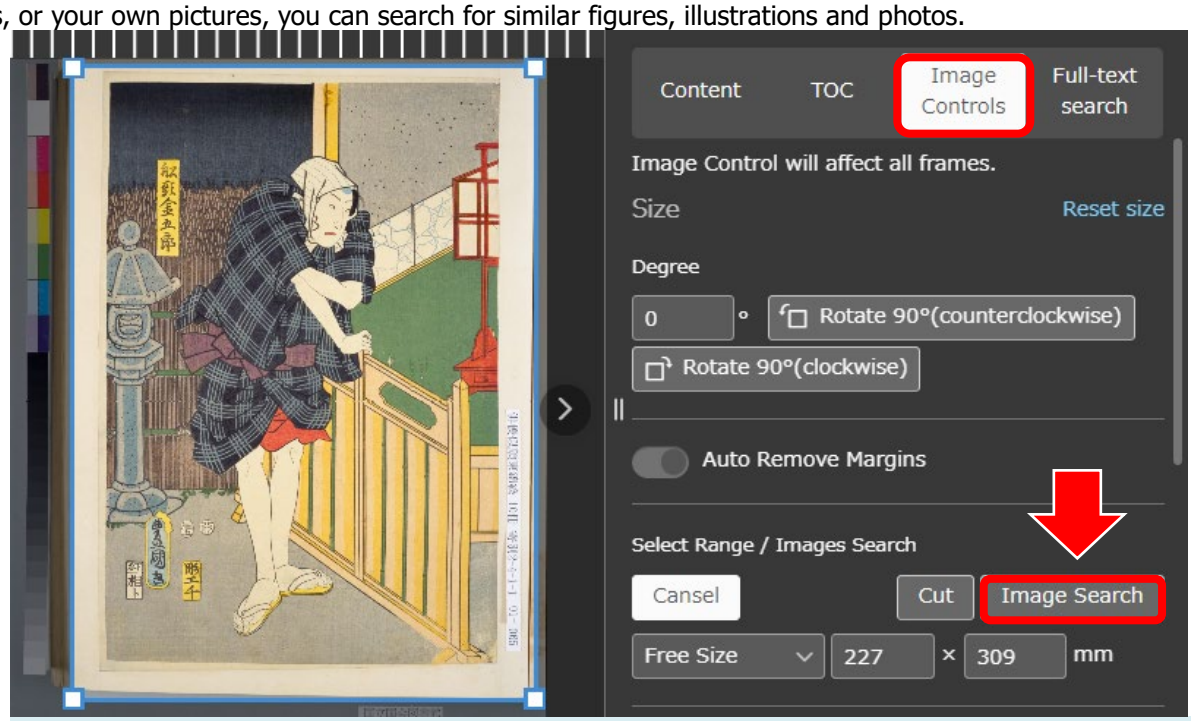
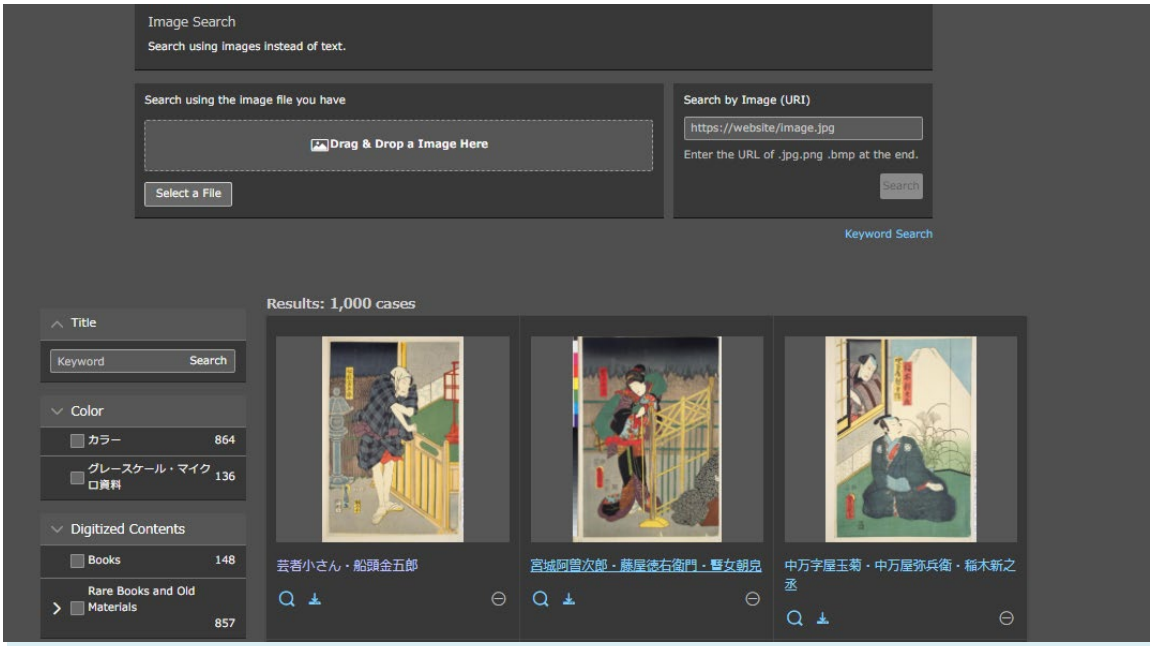


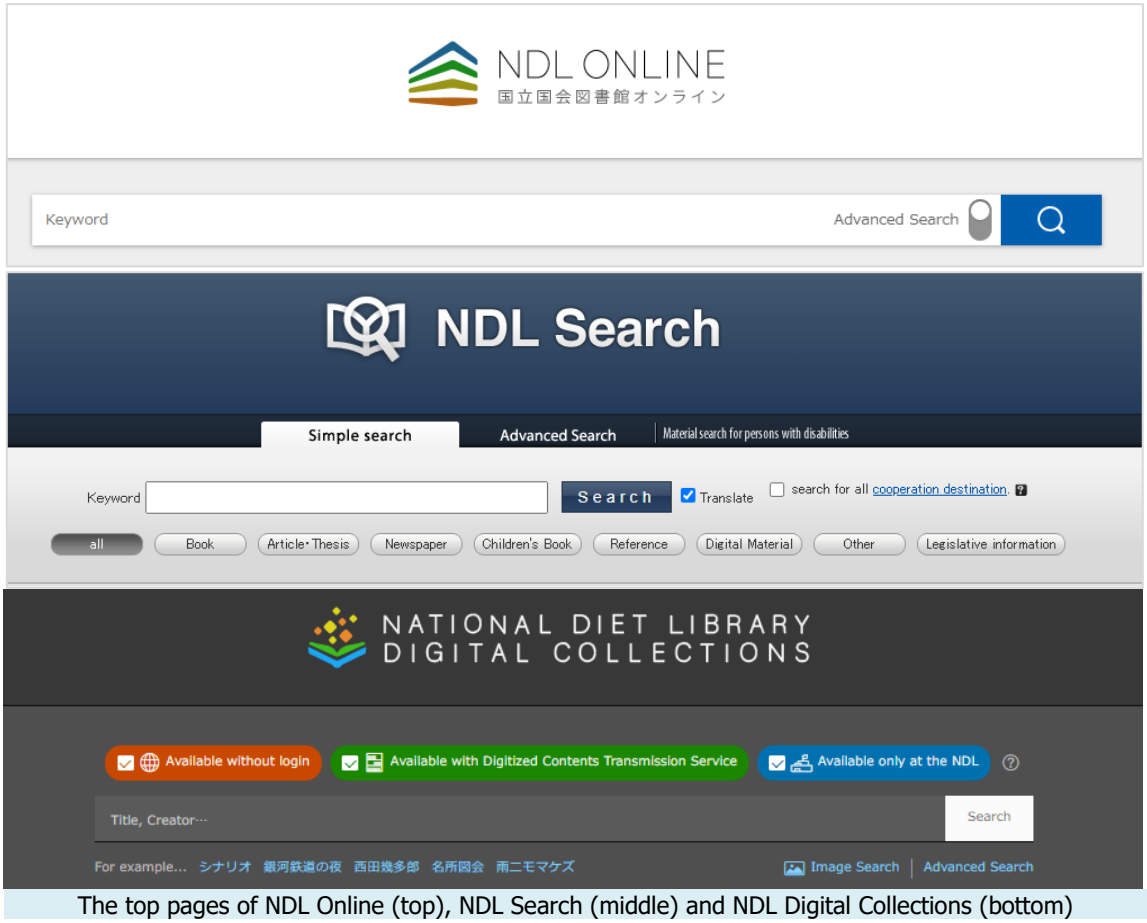
Image search from the browsing page
You can select the range of the picture and search similar pictures from the "Image Search" button.



The results of an image search

4. Single Sign-On on NDL Services

Single Sign-On is a system that allows you to use multiple web services with a single authentication. When you log in to one of NDL Online, NDL Search and NDL Digital Collections, you will be logged in to and can use all three websites. This enables you to use the library services more smoothly, such as applying for the remote photoduplication service for materials in the NDL Digital Collections.



The top pages of NDL Online (top), NDL Search (middle) and NDL Digital Collections (bottom)

(Translated by MACHIDA Sumika)

Events

2022 Online Exchange Program between the National Diet Library and the National Library of China

The 2022 Online Exchange Program between the National Diet Library (NDL) and the National Library of China (NLC) was held online on November 30, 2022.

The program is shown in the table below. Detailed presentations were made in each session, followed by lively discussions during the Q&A session.

Directors' meeting	Mr. YOSHINAGA Motonobu, director general of the NDL Mr. XIONG Yuanming, director of the NLC
Session	Theme: Development and construction of smart libraries • Expansion of Library Services and Future Prospects Enabled by Digitization of Materials and OCR Full-Text Conversion Mr. KIMEZAWA Tsukasa, deputy director general of the Digital Information Department, NDL • Knowledge Project of Digital Resources in the National Library of China Mr. XIE Qiang, director of Digital Resource & Service Department and director of Information Technology Department, NLC

Program

Mr. Kimezawa reported on digitization and OCR full-text conversion initiatives for the NDL collections, as well as expanded library services made possible by these projects, such as Digitized Contents Transmission Service for Individuals, full-text search service, and new experimental services such as the NDL Ngram Viewer, which utilize full-text data of digitized materials.

Mr. Xie reported on the NLC's knowledge project of local documents published during the Minguo period and others, including the extraction of knowledge data on people, organizations, events, and others from full-text data or data visualization such as the construction of knowledge maps.

NLC and NDL participants

Reference:

- Chronology of Mutual Visit Programs between the National Diet Library and the National Library of China

Related articles from the NDL Newsletter:

- Reports of past programs in the NDL Newsletter

Selected list of articles from NDL periodicals

The NDL Monthly Bulletin
No. 741, January 2023/ No.742, February 2023

If you click the volume number of each issue, you can read the full-text of NDL Monthly Bulletin. The text is provided in PDF format (in Japanese).

No. 741, January 2023 (PDF: 7.55 MB)

- New Year Greetings for 2023
- <Book of the month — from NDL collections>
  Kokkai no hanashi: For the next generation
- A library in the National Diet Building
- The National Diet Library's Detached Library in the Diet
- An interview with HORIUCHI Masaaki:
  Asking a professional of architectural history
- Where was the library in the Japanese Imperial Diet Building?
- Expansion of the acquisition of e-books and e-journals: Introduction of the e-legal deposit system for online publications
- Exhibition commemorating the 20th anniversary of the Kansai-kan
  World Expo Time Travel
- <Tidbits of information on NDL>
  For whom does the Parliamentary Documents and Official Publications Room exist
- <Books not commercially available>
  Kakugi fugi jiko no kenmei to mokuroku
- <NDL Topics>

No.742, February 2023 (PDF: 4.85MB)

- <Book of the month — from NDL collections>
  Potteries and porcelains depicted in the Meiji period: *Tojiki isho hyohon*
- Reading HAYASHI Hikaru's music manuscripts
  YOSHIDA Muneaki
- Travel writing on Japanese libraries (special edition)
  Kyoto Institute, Library and Archives
- Working at the Kyoto Institute, Library and Archives
- Tidbits of acquisition at the Kyoto Institute, Library and Archives: From various donated materials
- The NDL in figures: From the Annual Report of the NDL, FY2021
- <Tidbits of information on NDL>
  Connecting history and will
- <Books not commercially available>
  Hyakko ryoran: Takasago korekushon kaori no zukan: Tokubetsuban
- <NDL Topics>