

Articles by NDL staff

Hackathon: Let's make maximum use of the NDL's data

Research and Development for Next-Generation Systems Office, Digital Information Distribution Division,
Digital Information Department

This article is a translation of the article in Japanese in the NDL Monthly Bulletin No. 660 (April 2016).



1. Introduction

On November 28 and 29, 2015, the National Diet Library (NDL) held "Hackathon: Let's make maximum use of the NDL's data" at the Tokyo Main Library. This article describes both the event itself and some of the prototype applications that were developed during the event.

Event schedule:

Day 1: Saturday, November 28, 2015

a.m.

- Overview of the NDL data and application programming interfaces (API)
- Introduction to the Ideathon: "Let's make maximum use of the NDL's web page"
- Team building

p.m.

- Studying ideas and prototype development
- Progress reports

Day 2: Sunday, November 29, 2015

- Prototype development
- Presentation of results (Demonstrations of prototypes, discussions, wrap-up session)

2. What is a hackathon?

The word "hackathon" is a portmanteau of the words hack and marathon. It refers to an event in which people from different backgrounds, such as programmers and web designers, cooperate in the development of prototype products—such as applications for smartphones or computers, web tools, or digital content—and work together intensively for a limited amount of time. Although this was the first event of its kind to be held at the NDL, we were delighted to have more participants than we initially expected.

3. Team building through presentation of ideas

The subject of the hackathon was utilization of data from the [National Diet Library](#), the [National Diet Library](#) (Japanese), the [National Diet Library](#) (Japanese), and open datasets from the [National Diet Library](#). At the beginning of Day 1, an NDL staff member provided an overview of the available data included in the above contents and the application programming interfaces

(API) used in acquiring and utilizing the data.

Subsequently, Professor Takashi Harada, NDL Adjunct Researcher from Doshisha University, reported on the results of an event called Ideathon " " , another NDL Open Data workshop that was held in February 2015. After that, the participants formed groups and discussed ideas for prototype applications. Ideas were presented freely and as many as ten good proposals were discussed. After that, each participant chose a project to work on and five teams of three to six people each were formed.

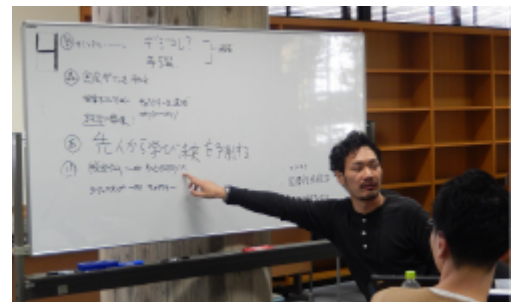


Image 1:
Discussing ideas



Image 2: Choosing projects

Image 3:
Presentation
of ideas



4. Go for Development!

Prototype development started on the afternoon of Day 1. Participants cemented their ideas by writing them down on a white board and actually started coding (writing a program) immediately on a computer to develop their prototypes. At the end of Day 1, each team made an interim report of their progress.

Even after the Day 1 session officially closed, some of the teams stayed at the venue and continued work until 9:00 p.m. to advance their development as far as possible in the limited time available. A few even continued to work at home until dawn. Although there was concern that some participants had worked a little too hard and looked tired the next morning, everyone was in high spirits and some teams shouted out as their prototypes steadily neared completion.

During the presentation of results at the end of Day 2, every team explained the background to their ideas and the techniques they used for development. Then they demonstrated their prototype and had a lively discussion afterwards.



Image 4: Studying ideas on the white board

5. Example of prototype: Digicolle Clipper

The word development is used frequently in this article; what are the actual steps of the development process during a Hackathon? Let's take a look at the development of the Digicolle Clipper prototype. This application was designed to facilitate the clipping and sharing of images from the NDL Digital Collections.

(1) The background of the idea

Much of the content available in the NDL Digital Collections was published a long time ago and might seem a little hard to appreciate. Looking through the database carefully, however, it is easy to find images and texts that are still intriguing. What's more, material for which copyright protection has expired may be used freely without applying for permission. The Digicolle Clipper team comprised four participants with an idea that would make it handier and easier than ever to find and share interesting images in the NDL Digital Collections.



Image 5: Pictures of the Idea

(2) Creation of development environment

Three of the four members were engineers with the know-how to make a core mechanism for opening applications. After the team was formed in the afternoon of Day 1, they created a development environment based on the ideas of the three engineers: to use software development tools that link programs on the Internet and to use free web services for sharing information.

(3) Mapping out a development environment

After creating a development environment, they next mapped out a development plan. Digicolle Clipper is intended for use in clipping and sharing images on SNS. As a corollary of their idea, the shared images are expected to attract additional page views for the NDL Digital Collections.

One of the members had already developed an application for clipping images from a web browser, and this became the base for development of the Digicolle Clipper. Using a white board, they schematized and studied the relations between an application to clip images, a server for storing clipped images, and metadata (bibliographic data) obtained from the NDL Digital Collections.

Consequently, they came to the conclusion that the prototype should have a function for sending the clipped images to a server and a mechanism for displaying the stored images in a list or other format.

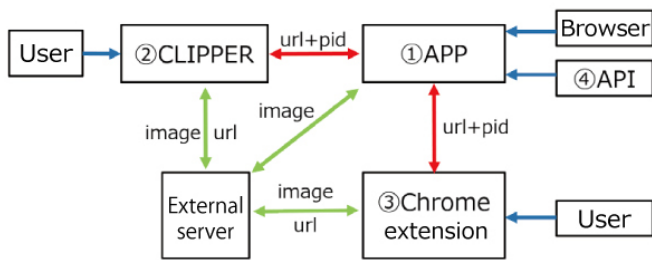


Image 6: Concept map of functionality

(4) Development of the application and extension functions

After each member was assigned an area of responsibility, the rest of the afternoon on Day 1 was spent developing the application body and expansion functions. On the morning of Day 2, they shared information on their progress from Day 1 and the challenges they had yet to overcome.

A function to share the clipped images on SNS was developed afterward. In the afternoon, they prepared a demonstration to be made during the presentation of results: They selected example images, clipped them with the application, and stored them in the server. In addition, they also developed a display function for smartphones.

(4) Demonstration

In their achievement presentation, the team demonstrated the display function, which was developed for use with devices such as PCs and smartphones.



Image 7: Functions of DigiColle Clipper

The presentation led to a lively discussion. Some participants were very keen to use DigiColle Clipper right away and others had suggestions for improving the image display. There were also suggestions from a librarian's perspective like clipping and storing images with bibliographic data included has the potential to trigger user interest in the original materials.

The team members discussed many functions that they wish to have, such as one for obtaining bibliographic data by making the application more specific to the NDL Digital

Collection, for categorizing clipped items and displaying them in order, and for specifying the number of items to display.

6. Interview with participants and their prototypes

(1) Mr. Naoki KOMADA - Nohon!-



Image 8: "Nohon!" (Legal-deposit publications)

By using bibliographic and holdings data of the NDL Search, this application enables users to verify whether or not a publication has been deposited with the NDL. You can also post the information on Twitter automatically.

Q: What brings you to this event?

A: I knew that the NDL has released APIs, but I had missed the opportunity to learn about them and had never used them before. And since this was also a good chance to develop an application, I thought I would participate.

Q: How was the Hackathon?

A: I learned quite a bit about the existing APIs through this event. I was also very pleased to have an impetus for learning about social conditions related to library collections.

Q: Thank you. Could you outline your prototype?

A: We developed a web application called "Nohon!" As the name suggests, this app was developed to verify whether or not a particular publication has been already deposited with the NDL. The NDL Search is useful for checking title, publisher and ISBN. But input errors sometimes lead to empty search results, so we can't always find things just by searching the database. Our idea was to make it possible to refer to collections of other public libraries as well as publication information. NDL Search is used to search information on the holdings of public libraries.



Image 9: Concept map of functionality

Q: What was the most interesting challenge you faced in your prototype development?

A: We did not have much trouble in the Hackathon, since we successfully developed an information acquisition function. If an API is not organized well, it is difficult to use in an application. I think the NDL offers user-friendly APIs. We did not use the Web NDL Authorities this time, but I was interested in the fact that SPARQL is available when running queries* in the Web NDL Authorities. Previous APIs obtain data mainly by putting search criteria into part of the URL. As for SPARQL, the grammar is defined, which I hope will be used more widely in the future, because that will help make APIs easier to use.

Q: If you have requests or expectations about data provided by the NDL, please let us know.

A: I hope that the NDL will continue to cooperate with other institutions and businesses by providing APIs to boost data and material coordination. I am looking forward to useful APIs maintained by the NDL.

*A query is a search request sent to a database or a system. SPARQL is a kind of computer language for such queries.

From the organizer:

Thank you very much for your interest in the Web NDL Authorities! We are happy to receive your requests and expectations. The NDL plans to continue maintaining APIs.

(2) Mr. Itsuro TAJIMA -DigiColle Clipper-

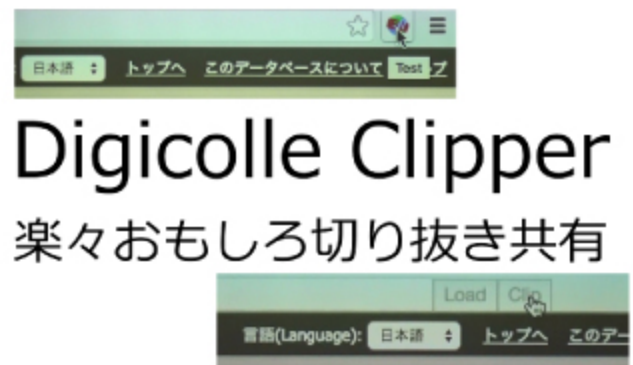


Image 10: DigiColle Clipper

This application facilitates the clipping and sharing of images from the NDL Digital Collection.

Q: What brings you to this event?

A: I have a friend who enjoys discovering intriguing items in the NDL Digital Collections, which stimulated my own interest in classics. But the archive contains such a huge amount of material that it is intimidating to get started. I thought it would be nice if we could create an interesting and easy-to-use application. So I was eager to participate in a Hackathon organized by the NDL.

Q: How was the Hackathon?

A: I enjoyed the Hackathon to the fullest as we developed a prototype and received a lot of ideas on it from other participants.

Q: Thank you. Could you outline your prototype?

A: We developed a service that facilitates the clipping and sharing of images of the Digital Collections. We created a PC application for clipping images, a viewer for PCs and smartphones, and set up a server for sharing the clippings. We also wanted to link bibliographic data to the clippings, but could not finish it in the limited amount of time we had.

Q: What was the most interesting challenge you faced in your prototype development?

A: We faced a number of unfamiliar things, like inter-process communication, which were difficult challenges. My teammate revised my program into more modern code to make it easy to maintain. I was surprised to see how neat the revised code was.

Q: If you have requests or expectations about data provided by the NDL, please let us know.

A: I would like the NDL to offer an API or other means for freely acquiring images that are in the public domain. I think that there would be some issues to overcome, including server load. But I believe that the NDL Digital Collections would become a very important source of image data if there were APIs for accessing the images more easily.

From the organizer:

Thank you so much for developing an ideal application for utilizing the NDL Digital Collections. We would like to continue expanding the scope of open data offered by the NDL.

7. Other prototypes developed in the Hackathon



Image 11: Tango Tango



Image 12: The Diet Poker

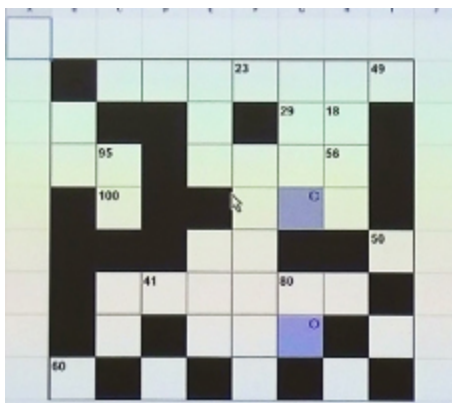


Image 13: Solve the mystery of libraries!

“Tango Tango”

An application that uses bibliographic data from the NDL Search to give feeling to *tango* (which means “words” in Japanese) with sound and diagrams.

“The Diet Poker”

A social game for providing an opportunity to learn more about Diet Members’ activities using data from the Full-text Database System for the Minutes of the Diet.

“Solve the mystery of libraries!”

An application that generates crossword puzzles on books featured in library events using bibliographic data from the NDL Search.

8. Conclusion

The Hackathon was a two-day event which required vitality and stamina from the participants, as they worked hard to create apps with only a limited amount of time for studying prototype ideas, designing, and developing their application. The participants cheered other teams’ demonstrations, posed questions, and gave advice to enhance the quality of the developed prototypes.

The participants enjoyed a stimulating and invigorating experience as they watched and then discussed the presentations of other teams, first from a users’ viewpoint and then from a developers’ standpoint. They shared a sense of accomplishment and enthusiasm for further improvement.

Participants’ comments included: “It was a good opportunity to master API,” “I became more familiar with the Full-text Database System for the Minutes of the Diet,” “I would like to join in this kind of event again when it is held,” and “It would be better if we had more time to exchange our opinions.”

This event provided an opportunity to meet and discuss programming with engineers and other people we ordinarily do not have a chance to meet face to face, and we found that NDL data is useful even outside of libraries. We will continue to promote use of a wide variety of NDL data.



Image 14: Best smiles