NII Today

91

Mar. 2021

National Institute of Informatics News

Expectations for NII RDC, NII's next-generation research data platform

Promoting open science and encouraging innovation

HASHIZUME, Atsushi
[Counselor for Information Technologies, Research Promotion Bureau, MEXT]

Roundtable **NII×Nagoya University**
NII RDC research data platform is finally set to launch!

YAMAJI, Kazutsuna
[Center Director, Research Center for Open Science and Data Platform, NII]

Research data management platform GakuNin RDM is now fully operational

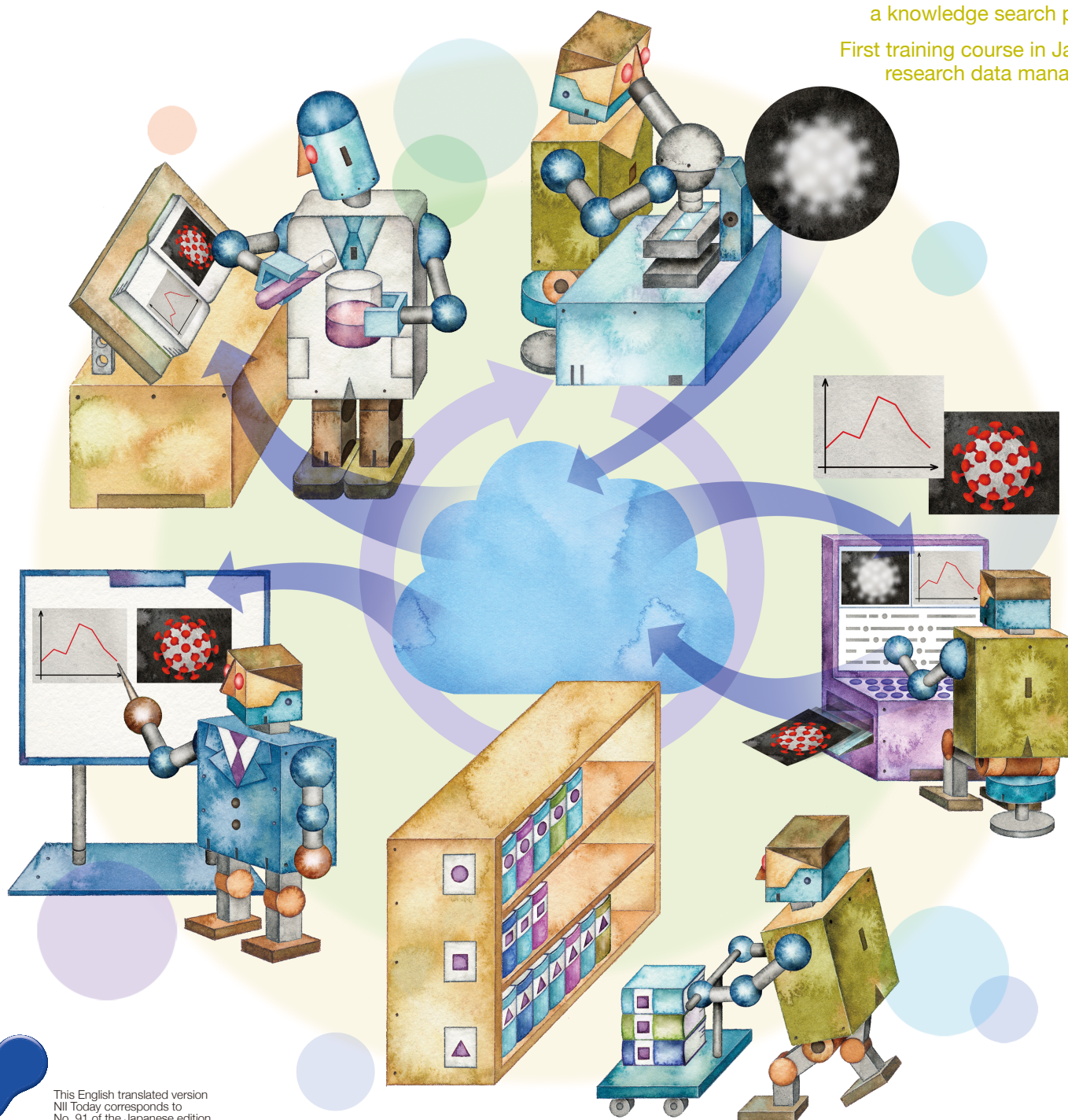WEKO3 supports data publication by Japanese academic institutions

CiNii Research's new role as a knowledge search platform

First training course in Japan for research data management

Feature

NII Research Data Cloud Set for Full-scale Launch

A research data platform to support open science

This English translated version NII Today corresponds to No. 91 of the Japanese edition

Expectations for NII RDC, NII's next-generation research data platform

Promoting open science and encouraging innovation

HASHIZUME, Atsushi

Counselor for Information Technologies, Research Promotion Bureau,
Ministry of Education, Culture, Sports, Science and Technology

Interviewer: TAKI, Jun-ichi

NIKKEI INC. Senior Writer

Japan's national "Integrated Innovation Strategy" and "Science and Technology Basic Plan" both stress the promotion of open science. The idea of returning research findings to society and the scientific community is about actively disclosing research data to foster cross-disciplinary research and promote innovation. Another aim is ensuring the transparency of research. The National Institute of Informatics (NII) is starting to offer a research data platform to facilitate effective storage, management, and use of the research data possessed by universities and other institutions. To find out more about the importance of developing this data platform, we spoke to HASHIZUME, Atsushi of the Ministry of Education, Culture, Sports, Science and Technology's (MEXT) Research Promotion Bureau, who currently serves as the government's Counselor for Information Technologies.

Revitalizing the research environment through open data and digital transformation

—— To begin, could you tell us about what is behind the current push to open science?

Hashizume: To promote science, it is increasingly important to share and publish research findings and data and to harness the power of a wide variety of people. The reason is that this

HASHIZUME, Atsushi

After graduating from Waseda University in 1994, he joined the (then) Science and Technology Agency. After a stint at the Ministry of Education, Culture, Sports, Science and Technology (MEXT) working on industry-academia collaboration, nuclear power damage compensation, and institutional reform relating to science, technology, and academia, he served as First Secretary at the Japanese Embassy in Canada and as a professor at Hiroshima University, before assuming his current position in April 2019.

allows us to make the best use of the fruits of research already obtained to advance our research efforts. This open access to data will not only be valuable to individual researchers, but it will also serve as a valuable resource for people investigating fields of study outside the realm of academia. This approach will also encourage interdisciplinary study and collaboration. And in terms of the so-called "fourth paradigm" of data science, I think this can contribute to the development of a new "data-intensive" science, distinct from the traditional research methods focused on observation, theory, and simulation.

I also believe that the disclosure of data is important from the viewpoint of "citizen science," because it empowers many people who are strongly interested in science despite not being professional researchers. On top of this, disclosure of data helps to ensure the transparency of research processes and the integrity of research.

—— In these past few years, Japan's research capability has slumped according to many observers. Is this initiative going to improve the situation?

Hashizume: Various factors have contributed to the decline in research capabilities, so open science alone is not going to be a simple solution. However, if we consider openness as one of the two wheels of a cart, along with digital transformation (DX), it can serve as a powerful revitalizing force for the research scene in Japan.

In an analysis by the National Institute of Science and Technology Policy (NISTEP) of MEXT, I noted that the proportion of research that is likely to produce seeds for novel research is lower in Japan than in other countries. There seems to be a tendency here for research to concentrate on established mainstream fields. This is fine, but to make advances in science, it is essential to try new things. And to promote this, we clearly need to share all kinds of data and findings, making them easily accessible to everyone. If we make it easier for people from diverse

fields to get involved in research and give them more opportunity to learn, we can enhance our research capabilities.

Shaping a new s style of research through digitalization

—— Japan's 6th Science and Technology Innovation Basic Plan emphasizes digital transformation

Hashizume: Information technology was a vital issue even before the outbreak of COVID-19, but with the pandemic, we have been forced to address it. The national government has established the new Digital Agency to deal comprehensively with social innovation. And to promote digitalization in research, education, sports, and culture, MEXT has also established its own the Headquarters for Promoting Digitalization headed by the minister, HAGIUDA, Koichi.

Digitalization does not mean simply digitizing existing information by converting it to 0s and 1s. It means enhancing the environment for research and education by developing new ways of doing things. One example is the promotion of new data-intensive science, as I mentioned earlier. At the same time, it is also important to make large research facilities and laboratories remote-ready and smart, so that even after the pandemic, researchers can access and utilize them without the traditional constraints of time and place.

These kinds of research environments can also contribute to "work style reform" by allowing researchers to work in different ways, according to their lifestyles. Automation will also make it easier to verify experiments and it can expand the possibilities for reproducing the specialized skills of the technicians who support experiments—cell culture techniques, for example—in various locations, without the limitations of affiliation or location.

—— Are other countries ahead of Japan in developing data infrastructure?

Hashizume: There are various movements, like the EU's effort to create a big data platform called the European Open Science Cloud. Also, at the G7 Science and Technology Ministers' Meeting in Tsukuba, Ibaraki in 2016, Japan shared its views with the other countries, even proposing a resolution to promote open science. Given these circumstances, I think Japan needs to accelerate the development of its data infrastructure.

Challenges: repository collaboration, making rules, and cultivating personnel

—— What are your expectations for the next-generation data platform that NII is developing?

Hashizume: Although storing and sharing data is well established in some fields, such as global environment studies, other fields have no repositories for storing data. It is important to create a platform that connects and supports all fields. The fact that an inter-university research institute like NII is launching a data platform service for the entire university community and that instead of competing with existing infrastructure it is aiming for gradual integration through metadata collaboration is very important to Japan as a whole, as a data infrastructure creation initiative. I have high expectations for it. At MEXT, we are very keen to work seriously on this project with NII.

—— What do you see as the main challenges in expanding data-intensive open science, on the other hand?

Hashizume: This may seem to contradict what I said earlier, but above all, I think it is important to set up repositories for storing data safely and securely. Such repositories should not be closed; they should be interconnectable. NII and MEXT will work together to realize this goal as part of the next-generation Science Information Network (SINET). This is the first challenge.

Secondly, when the platform is ready to go, we will need to get researchers in different fields to collaborate with IT researchers to promote new science through data use in all fields. Formulating rules is also essential. How will we manage our research data? We need to think about open or closed strategies; should data be open or shared only with the parties involved? There are also more and more opportunities for using data generated from social activities for research, including personal data. But we also need to think about rules for the handling of such data.

Training the people to manage data is another essential need. Rather than each research institute promoting data science in isolation, we need to have people who can look after the whole platform by linking everything together, like sharing good practices and offering support to resolve any problems.

NII is ideally positioned for this role. On top of its core function of overseeing the data platform, NII is equipped with strong research capabilities. Since it offers research platform services, such as SINET, as well as handling cutting-edge data science research, NII can feed back the needs of the research frontlines to data science researchers and make use of them. I would like to see NII roll out a variety of services on top of its research platform and demonstrate the possibilities of the new era by creating advanced case studies of what can be achieved.

—— The transition to digital will eventually be a good thing if we stick with it, but it will probably be hard for researchers initially. What would you say to researchers on the frontlines?

Hashizume: You often hear people say things like, "make the best out of a bad situation," but it is not easy to address a new situation when there are so many challenges. I would like to show them what the future will look like after we have gone ahead with digitalization and open science. The same with formulating rules; rather than imposing regulations, I would like to share the knowledge that data can be effectively managed in this or that way, so that the researchers can relax a little more in their work.

A Word from the Interviewer

Hardly a day goes by that I do not come across the expressions "digital transformation" (DX) or "digitalization." The COVID-19 pandemic has made us painfully aware of how slow Japan's society and government has been to digitalize. While I do not welcome every aspect of digitalization with open arms, I feel that Japan needs to prepare itself to inevitably move with the times and even get ahead of global trends.

As Mr. Hashizume pointed out, the NII is a uniquely placed Japanese research institute. It both provides research information infrastructure, as typified by SINET, and also operates as a research institute equipped with fundamental information science capabilities. A clear example demonstrating this function is the series of cyber symposiums on online education that NII held during the pandemic. Sharing information in this way helped many universities that were pressured to switch to online education delivery. It would be great to see NII do something similar for its data platform.

TAKI, Jun-ichi

Senior Writer, NIKKEI INC.
Born in 1956. After graduating from the School of Political Science and Economics at Waseda University, joined Nikkei, Inc. After working in branch offices and covering corporate news, began covering science and technology, as well as environmental fields, starting from the mid1980s. Authored "Eco-Uma ni Nore!" (Shogakukan) and co-authored "Kansensho Retto" (NIKKEI, Inc.), among others.

Photography by SATO, Yusuke

NII×Nagoya University
NII RDC research data platform is finally set to launch!

Research data management challenges and prospects come to light

AOKI, Takaaki

Professor, Information Strategy Office, Information and Communications, Nagoya University; Chair, Research Data Management Subcommittee, Academic eXchange for Information Environment and Strategy (AXIES)
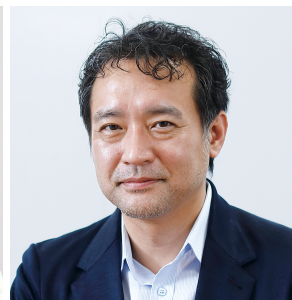
HAYASHI, Kazuhiro

Special Advisor, Information Management Section, Nagoya University Library ; JAIRO Cloud team member, Content Distribution Promotion Working Group, Japan Consortium for Open Access Repository (JPCOAR)

TANAKA, Sachie

Manager, East Area Library Management Group, Higashiyama Campus Library Section, Nagoya University Library

YAMAJI, Kazutsuna

Professor, Digital Content and Media Sciences Research Division, NII; Center Director, Research Center for Open Science and Data Platform (RCOS)

In FY2021, NII Research Data Cloud (NII RDC) began full-scale operation. NII RDC is made up of three platforms: a research data management platform (GakuNin RDM), a publication platform (JAIRO Cloud), and a search platform (CiNii Research). In 2020, the management platform was operated on a trial basis in preparation for the launch of the new service. In this context, Nagoya University, which has been a leader in this field, has worked on improving the environment for research data management, focusing on its libraries and Information Technology Center. We invited three of the people from Nagoya University who have been involved in preparing the RDM system, formulating rules, and developing and supporting the data publication platform, together with Prof. YAMAJI, Kazutsuna, Center Director of NII's Research Center for Open Science and Data Platform (RCOS), to talk about their current efforts. They spoke about the challenges they have faced and their outlook for RDM.

A sense of crisis at universities relating to research data management

——— **Full-scale operation of NII Research Data Cloud (NII RDC) began in FY2021. How is this service different from previous ones? And why is it important?**

Yamaji: Until now, NII mainly offered services relating to published papers, like "JAIRO Cloud," an institutional repository service for storing, operating, and publishing academic articles, materials, and research findings in digital formats, and "CiNii," a literature search service. As well as advancing these services, NII RDC has launched a new research data management platform service called GakuNin RDM (Research Data Management). GakuNin RDM enables the storage and management of the research data and related materials generated in the process of academic research activities. (See Fig. 1.) This is a totally new service that is directly connected to research activities. For NII, it has been a big challenge.

The lifecycle of an open science process is outlined in Fig. 1. Firstly, you formulate a research plan, then collect data through experiments and other means, and analyze the data. Finally, you publish the findings in an article or other document. This stimulates ideas in other researchers and leads to new research. The biggest difference now is that the closed part on the right-hand side of the cycle—that is, the research data that were previously not disclosed to the public—is now managed on a shared platform. By linking this to the publication platform, any research data that researchers have the discretion to disclose can be published (made public) for the purpose of furthering the development of science.

Previously at universities, the work of publishing papers through institutional repositories and other means was mainly handled by libraries. Now that this task has been expanded to include the management of research data, universities face a greater challenge.

Aoki: The reality is that researchers and laboratories all manage

research data in their own particular way, in a process that is very much closed. It is impossible for an outside observer to know what kind of research data each researcher is handling and how they are managing the data. Every research field also seems to have its own way of managing data. For researchers, research data are assets, but with the abolition of the chair system and the expansion of the special appointment system, it is not uncommon for data obtained at laboratories, research notes, and know-how to be entirely lost when a faculty member retires or transfers. If this situation persists, universities risk losing more and more valuable assets, so they are feeling a growing sense of crisis. Such losses are a loss for the whole of academia, so I think it is vital to preserve and manage research data efficiently through digitalization.

Yamaji: Losses of research data due to damaged hard disks and to people forgetting where the data were stored are occurring all over Japan. Of course, if researchers all had to buy their own data storage drives or services to manage their data, it would be laborious and costly for them. Research integrity is another issue. Current guidelines stipulate that, as a rule, research materials should be kept for at least 10 years, but this requirement is not being fully satisfied. If universities can actively set up an environment in which research materials and data can be managed by means of an institution-wide system, researchers would benefit significantly. I believe that there is a deep-seated need for this kind of solution.

A new role for university libraries

——————— What kind of initiatives has Nagoya University taken?

Aoki: In early 2019, the president of the university instructed us directly to promote research data management and open science, so an investigation system was quickly set up. We first formulated an academic data policy to define a course of action, making clear that the university's mission is to provide an environment that supports the management, publication, and utilization of research data. Faculty members connected with the university's library and IT infrastructure are leading the effort to develop the environment. (See Fig. 2.)

Hayashi: The library people are working mainly on the publication platform. Within the research data cycle, publication plays the essential role of circulating academic information and stimulating further new research. Up to now, we have stored and disseminated academic papers and other information using JAIRO Cloud, NII's institutional repository service. Currently, we are in the process of migrating to a new version of JAIRO Cloud (software name: WEKO3), which enables the publication of research data, as well as papers.

Yamaji: The system architecture of WEKO3 makes it easy to add new functions. The new system is just starting to be used, but we are planning to expand its functionality to make it even easier to use as we get feedback from the various institutions that have adopted it.

Hayashi: For libraries, research data present a new challenge because they have not dealt with such data before, but the ability to share issues and solutions with other academic institutions over the cloud using NII RDC is reassuring to them. If we can initiate a chain of events in which the early adopters become a model for further deployment, the system will spread rapidly. I expect that this process will lead to the establishment of a management system for academic information throughout Japan, thereby promoting greater dissemination of information.

Tanaka: If data are stored on the management platform from the beginning of the research process, it becomes easy to transfer the data to the publication platform. These benefits of system integration are not promoted enough, so I hope that libraries can not only support the publication process but also help with raising awareness of the benefits of research data management. As a library, we want to contribute to the global trend of making research data more open.
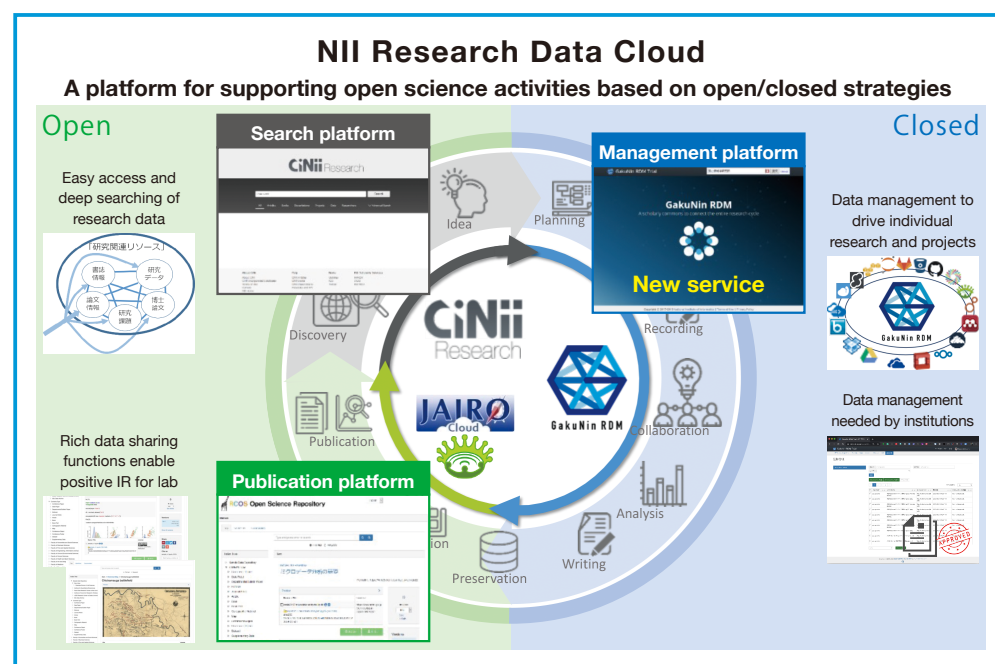

Fig. 1: NII Research Data Cloud
Within the lifecycle of research data, GakuNin RDM supports data management, JAIRO Cloud supports publication, and CiNii Research supports search. These offer a stress-free research data management system, with smooth data transfers between each service.

AOKI, Takaaki
Since April 2020, he has been involved in planning a university-wide information platform as a professor in the Information Strategy Office, Information and Communications, Nagoya University. Since his previous position at the Institute for Information Management and Communication, Kyoto University, he has worked on the subject of research data management with a view to establishing an information platform for supporting research.

HAYASHI, Kazuhiro
In 2018, he took up an appointment at Nagoya University Library, where he has been involved in repository operation and management. He has also participated in JPCOAR activities, specifically on trial operation of the next-generation JAIRO Cloud and on support for migrating institutions.

TANAKA, Sachie
Started working at Nagoya University as a library staff member in April 2008. As a member of the Training Sub-Working Group of the Open Science Research Data Platform Working Group of the Organization for Science Network Operations and Coordination in FY2020, she investigated continuous creation and updating of teaching materials and an operation system for personnel training related to the promotion of research data management.

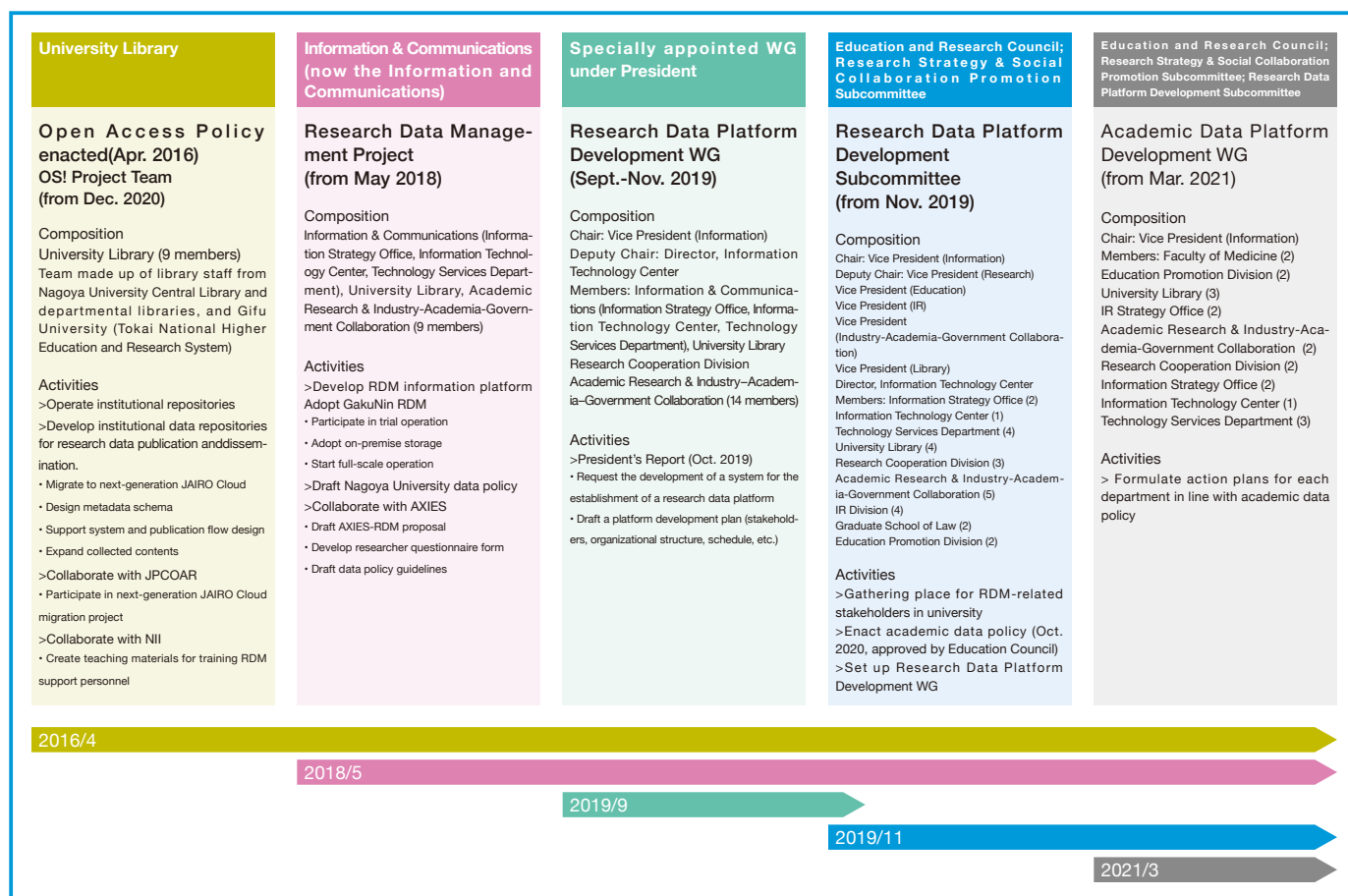| University Library	Information & Communications (now the Information and Communications)	Specially appointed WG under President	Education and Research Council; Research Strategy & Social Collaboration Promotion Subcommittee	Education and Research Council; Research Strategy & Social Collaboration Promotion Subcommittee; Research Data Platform Development Subcommittee
Open Access Policy enacted(Apr. 2016) OS! Project Team (from Dec. 2020)	**Research Data Management Project (from May 2018)**	**Research Data Platform Development WG (Sept.-Nov. 2019)**	**Research Data Platform Development Subcommittee (from Nov. 2019)**	**Academic Data Platform Development WG (from Mar. 2021)**
Composition University Library (9 members) Team made up of library staff from Nagoya University Central Library and departmental libraries, and Gifu University (Tokai National Higher Education and Research System) Activities >Operate institutional repositories >Develop institutional data repositories for research data publication and dissemination. • Migrate to next-generation JAIRO Cloud • Design metadata schema • Support system and publication flow design • Expand collected contents >Collaborate with JPCOAR • Participate in next-generation JAIRO Cloud migration project >Collaborate with NII • Create teaching materials for training RDM support personnel	Composition Information & Communications (Information Strategy Office, Information Technology Center, Technology Services Department), University Library, Academic Research & Industry-Academia-Government Collaboration (9 members) Activities >Develop RDM information platform Adopt GakuNin RDM • Participate in trial operation • Adopt on-premise storage • Start full-scale operation >Draft Nagoya University data policy >Collaborate with AXIES • Draft AXIES-RDM proposal • Develop researcher questionnaire form • Draft data policy guidelines	Composition Chair: Vice President (Information) Deputy Chair: Director, Information Technology Center Members: Information & Communications (Information Strategy Office, Information Technology Center, Technology Services Department), University Library Research Cooperation Division Academic Research & Industry–Academia–Government Collaboration (14 members) Activities >President's Report (Oct. 2019) • Request the development of a system for the establishment of a research data platform • Draft a platform development plan (stakeholders, organizational structure, schedule, etc.)	Composition Chair: Vice President (Information) Deputy Chair: Vice President (Research) Vice President (Education) Vice President (IR) Vice President (Industry-Academia-Government Collaboration) Vice President (Library) Director, Information Technology Center Members: Information Strategy Office (2) Information Technology Center (1) Technology Services Department (4) University Library (4) Research Cooperation Division (3) Academic Research & Industry-Academia-Government Collaboration (5) IR Division (4) Graduate School of Law (2) Education Promotion Division (2) Activities >Gathering place for RDM-related stakeholders in university >Enact academic data policy (Oct. 2020, approved by Education Council) >Set up Research Data Platform Development WG	Composition Chair: Vice President (Information) Members: Faculty of Medicine (2) Education Promotion Division (2) University Library (3) IR Strategy Office (2) Academic Research & Industry-Academia-Government Collaboration (2) Research Cooperation Division (2) Information Strategy Office (2) Information Technology Center (1) Technology Services Department (3) Activities > Formulate action plans for each department in line with academic data policy
2016/4	2018/5	2019/9	2019/11	2021/3

Fig. 2: Establishment of RDM support system at Nagoya University
For more information on the efforts at Nagoya University, see "Development of Promotion Organization of Research Data Infrastructure at Nagoya University," by TAKEYA, Kimie, as well as the 3rd SPARC Japan Seminar 2019 (https://www.nii.ac.jp/sparc/en/event/2019/20200207en.html).

Who supports research data management and how?

——— **To implement research data management, support personnel are indispensable.**

Tanaka: A working group of the Japan Consortium for Open Access Repository (JPCOAR) has led the creation of several educational resources for training research data management support personnel. Now, to get research institutions to make more use of these materials, the Training Sub-Working Group (TSWG) of the Open Science Research Data Platform Working Group formed at NII is preparing a list of the key tasks involved in research data management, along with the skills required to perform them. This skills list is designed to help research institutions with the selection, development, and evaluation of educational materials and programs for training support personnel. We have also added information on the professional categories capable of performing these tasks and the types of people who should be trained for them. So, we believe that this information can serve as a useful guide for building a systematic support and collaboration system suited to the circumstances of the particular institution.

Yamaji: The existing teaching materials were originally created based on overseas case studies and materials, so the TSWG has adapted them to suit the employment system and organizational structure of universities in Japan.

Tanaka: My overall impression of librarians is that many of them would be interested in research data management. However, since it is not directly related to their existing work, they do not know how to get involved in it. I really want to provide content that gives them an opportunity to take a first step toward research data management. Frankly, I had some doubts about whether research data management was really a task that libraries should take on, but when I analyzed all the tasks involved, I realized that they were not totally new. In fact, many of them could be done as an extension of or on top of work that libraries already do, like assigning metadata to research data registered in the publication platform.

Hayashi: Although there are some differences, such as the higher diversity of data formats handled and the cloud-based nature of the work, the role of libraries remains basically the same—that is, organizing and making available academic journals. Or in other words, they support educational and research activities through the distribution and storage of academic information. Once I realized this, I started to think about the best ways of delivering information to users, which was satisfying and interesting.

Aoki: Surprisingly, many researchers and students do not understand what the libraries and the Information Technology Center do. So, getting involved in research data management may be a good way for library and information technology center staff to work more closely with researchers and build relationships of mutual trust with them.

Tanaka: As Prof. Aoki says, it would be ideal if we could make researchers and students aware of the potential of libraries and other resources so that they can make use of them effectively when they want to publish their research data. Within the administrative structure of universities, the work of libraries has

been difficult to see from the outside, a factor that has limited its appeal. The work involved in research data management is so diverse that no single department can support it completely. Inevitably, collaboration with other departments (apart from the library) is necessary. By sharing capabilities in this way, it is possible to offer systematic support. I would be very pleased if we could improve the standing of the library through our involvement in research data support.

First, uncover the needs

———— Full-scale operation of the GakuNin RDM management platform commenced on February 15, 2021. How is the usage situation?

Aoki: Nagoya University was participating in the trial operation since last year. We used storage prepared by the university and connected it to GakuNin RDM. In other words, data remained on the Nagoya University system, but we used GakuNin RDM as a portal for data sharing. This is how we plan to continue using the service. We have not used GakuNin RDM for any large-scale research projects yet, but I have heard of a few cases of research laboratories using it.

From the examples I have seen so far, the appeal of GakuNin RDM seems to lie not just in managing research data within a research laboratory or university, but also in its use as a portal for facilitating collaboration with other academic institutions. Although it does not involve Nagoya University, I know of one instance of how GakuNin RDM has been used as a portal in an inter-university joint research project for creating a system to manage research data.

———— What kinds of issues have you come across?

Aoki: Since many researchers already have their own methods of managing research data, getting a new system established is no easy task. As suggested by the old saying, "You cannot put new wine into old bottles," I believe that the first step is to focus on literacy education for young researchers just starting out, encouraging them to use the system on totally new research projects. This will lead to the efficient expansion of system use. To spread this kind of new system, it is also essential to offer strong support to help people overcome their unfamiliarity.

Yamaji: Whenever you do something new, there will always be people who oppose it. So, this is a major challenge. It is extremely difficult to change people's thinking and behavior patterns. I have to say, though, that uncovering their needs and gradually changing their current workflow is the most interesting part of the project.

Aoki: Indeed, at Nagoya University, we are uncovering needs as we create our system. The needs undeniably exist, so the system definitely offers researchers greater convenience. If you want to manage your research data in your own way, you need to decide how to name files, where to save them, what items of information to record in your research notebook, and so on. As the use of GakuNin RDM gets established, its functions will expand and evolve, and its data management methods will become standardized.

All these things will significantly reduce the stress on researchers.

Yamaji: Recently, we have seen a growing number of cases of retiring professors who want to preserve their research data for the sake of academic development. At the same time, more and more often researchers need to publish their research data together with their papers. In the area of publication, we will identify these kinds of needs to gradually expand the service and change workflows. If we look ahead to the future with an eye on overseas trends, I expect that RDM will become an indispensable element of a researcher's work.

The ability to implement functions while listening to feedback from the university

———— What needs to be done to promote use of the system in the coming years?

Aoki: For research data management, merely sharing storage is not enough, because in that case you are just competing with various other cloud services. We need to go one step further than storage sharing by making the system more functional as a portal for research projects.

Yamaji: Researchers are very critical about the tools and systems they use every day. We need to create something that is easy to use and useful so that people do not say, "I cannot use this." We will be getting feedback on what skills are needed for research data management and what functions are needed from the experience of universities. And NII will be working with universities to improve and refine the system continually through the implementation of a PDCA cycle. We are starting on this now, so we will be entering an even more important phase.

Aoki: We need to make our system of cooperation better than ever. In terms of functions, I think the most important thing is ease of use. If there are too many things to do as preliminaries or preparations, people will tend to give up before even using it. If the university participates in the Academic Access Management Federation (GakuNin), the GakuNin RDM system can be used immediately by just linking with an identity provider (IdP). However, since the operation of IdPs is a problem for institutions, the hurdle is very high for end-user researchers. I would like to see more and more institutions joining GakuNin so that more and more colleagues can easily make use of GakuNin RDM.

Tanaka: Even from the perspective of supporting research data management, if the system is not easy to use, few people will be able to teach it, so use of the system will remain very limited. Thus, we not only want to use the system that NII has built, but as a university, we want to actively channel user feedback and case study information back to developers, to help make the product better.

Hayashi: There are still very few examples of publication platforms in Japan, so if the NII RDC can serve as a hub for sharing case studies with other institutions and increasing the number of colleagues, I think that this can provide a foothold for further development.

Yamaji: Our greatest strength is our ability to listen to what universities say and implement the functions they need. This is because our relationship of trust with universities is a vital asset for NII. In the process of building the NII RDC, we have established a system that allows us to have conversations with university people and young academics on an equal footing. For young people, tackling big challenges together across sectors can be a very good experience. If we succeed with this, I look forward to building stronger relationship with them.

Written by Hata, Chisato

Research data management platform GakuNin RDM is now fully operational

Supporting advanced research data management

On February 15, 2021, the GakuNin RDM (Research Data Management) service for supporting universities and research institutes in managing their research data commenced full-scale operation. The service, developed by the Research Center for Open Science and Data Platform (RCOS), is the first research data management platform running on the NII Research Data Cloud (NII RDC), a research data platform designed by NII to promote open science. What functions does it offer as a platform for RDM, which is the key to open science practice?

KOMIYAMA, Yusuke
Associate Professor, Digital Content and Media Sciences Research Division; Research Center for Open Science and Data Platform, NII

NIITSUMA, Akira
Head, Research Data Cloud Team, Scholarly and Academic Information Division, Cyber Science Infrastructure Development Department, NII

Service aims at "research promotion" and "research integrity"

On top of implementing advanced data management in large-scale research, in recent years, organizing and structuring data for storage and management have become absolute requirements of pursuing research. For example, it is now obligatory to include such data management details in plans submitted to research funding organizations. To ensure that these kinds of research data management duties are appropriately executed, universities and research institutions are starting to take over these duties from individual researchers

To support this change in the academic world, NII has launched GakuNin RDM, the first research data management platform service running on the NII RDC. The service's main purpose is to offer IT infrastructure for managing and sharing research data, with the two key goals of "research promotion" and "research integrity." (See the Fig.) And as Professor KOMIYAMA, who led the development team, explained, it was essential to make the service easy to use for researchers and to provide institutions with a robust research data management system.

Firstly, from the viewpoint of research promotion, any institution participating in the NII GakuNin Academic Access Management Federation can immediately receive the service. The GakuNin RDM service can also be linked to external clouds and repositories, allowing previously accumulated research data to be imported. It is also equipped with data sharing and other essential functions for collaborative research. These functions are set up to allow easy use by individual researchers and institutions. At the same time, the service constantly records the trail of research data, thereby enabling follow-up investigation of research data in the event of research misconduct or suspicious activities. This feature serves as a deterrent to research miscon-

duct and helps institutions to manage their data securely.

Supporting safe data sharing and industry-academia collaboration

All the currently available functions of the service were trialed by 22 participating institutions in a verification experiment running from April 2019 to October 2020. While the Institute for Quantitative Biosciences at The University of Tokyo investigated methods of utilizing the research integrity functions of the service at the institute, Hokkaido University planned to link the data management platform to experimental and measurement devices. Like this, different institutions take their own approaches to RDM adoption. This real-world use of GakuNin RDM clearly revealed that the service can address a wide range of needs. For example, it can serve as a system for appropriately managing all papers, images, and raw data published by researchers, it can be linked with platforms for analyzing data imported from experimental devices, and it can be used for RDM applications for storage provided for internal use within an institution. Professor KOMIYAMA relates, "Especially, we receive many requests to connect the service to JAIRO Cloud, the data publication platform of the NII RDC, and to allow access to companies for industry–academia collaborations." Many institutions want to use GakuNin RDM not only for storage but in the future for data utilization and large-scale joint research too. Note that connection of the service to JAIRO Cloud is expected by the end of FY2021. This would both enable smooth data publication and also create a secure research data platform for preventing accidental data disclosures by passing the data transfer procedure from GakuNin RDM to JAIRO Cloud.

Full-scale operation of GakuNin RDM commenced on February 15. With this launch, operation of the service was transferred to the Scholarly and Academic Information Division. NIITSUMA, Akira, who handles the organizational and legal preparations of the service, explains, "The Scholarly and Academic Information Division has set up a Research Data Platform Steering Committee made up of around 10 of the participating institutions. In consultation with the steering committee, we will now work together with RCOS on ensuring continuing stable operation, as well as making functional enhancements and improvements." Although the application of service fees may be considered in the future to continue the service, for now the service is provided free of charge. GakuNin RDM, which has the power to revitalize research in Japan, will continue to serve up a comprehensive range of services running on its robust platform.
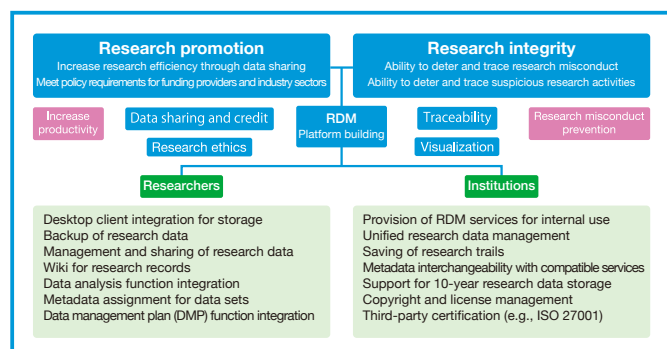
Interview/Report by IKEDA, Akiko   Photography by FURUSUE, Takuya

Fig: Overview of GakuNin RDM service. It supports researchers through "research promotion" and provides secure data management to institutions by ensuring "research integrity."

WEKO3 supports data publication by Japanese academic institutions

WEKO3, designed to run on JAIRO Cloud, has launched

Japan happens to be the second most advanced country after the U.S. in terms of institutional repositories, boasting a total of 876 universities and academic institutions with their own institutional repositories (as of Jan. 2021). Some 70% of these repositories utilize JAIRO Cloud, a cloud service for institutional repositories operated jointly by NII and the Japan Consortium for Open Access Repository (JPCOAR). JAIRO Cloud is currently being migrated and upgraded to the next version together with WEKO, the software it is built on.

HAYASHI, Masaharu
Project Assistant Professor, Research Center for Open Science and Data Platform, NII

HAYASHI, Yutaka
Head, Research Data Cloud Team, Scholarly and Academic Information Division, Cyber Science Infrastructure Development Department, NII

Revamping JAIRO Cloud as a research data repository

Institutional repository is a generic term for a series of services used by universities and academic institutions to store, manage, and publish papers, academic materials, and research findings in digital formats. In terms of the number of its institutional repositories, Japan is a world leader, but few academic institutions in the country are capable of building and operating their own repository. To support institutions, therefore, in 2012 NII rolled out its JAIRO Cloud managed service, which offers institutional repository functions. In the first year, 73 institutions adopted the service; currently, 638 are using it (as of Jan. 2021). The service is operated jointly by NII and the Japan Consortium for Open Access Repository (JPCOAR), made up of universities and other institutions.

HAYASHI, Yutaka, who oversees operation of the service says, "Ever since we launched JAIRO Cloud, we have been extending its functionality, incorporating the feedback and requests of numerous user institutions and the library community. So, in a way, the cloud service has grown together with its community."

In addition to responding continuously to the needs of end users, like universities and academic institutions, there is another role that the academic community expects JAIRO Cloud to play. Across the world in the past few years, there have been calls for more open access to research data, as well as research literature. Responding to this trend, NII decided to enhance the functionality of WEKO2, the software that powers JAIRO Cloud, to deliver the requested functions of literature repository and research data repository as a cloud service with an assured level of availability, reliability, and maintainability. Thus, NII set out to upgrade JAIRO Cloud to serve as a new institutional repository service built on WEKO3, which boasts several cutting-edge architectural features.

Professor HAYASHI, Masaharu, who worked on the development of WEKO3, says, "WEKO2 evolved into a refined literature repository service that catered attentively to its users' needs. Although WEKO3 will maintain all these capabilities, it will be built and operated as a new distribution platform for academic information that also serves as a research data repository.

WEKO3 is almost a complete rewrite of WEKO2, with a very different infrastructural configuration. Making use of a Kubernetes cluster running on a public cloud, PostgreSQL for its master database, RabbitMQ as a message broker, and Elasticsearch as a full-text search engine, WEKO3 boasts a modern architecture optimized for performance, load balancing, and automated operation. WEKO3 is based on the "Invenio 3" repository software framework developed by the European Organization for Nuclear Research (CERN). Professor HAYASHI (Masaharu) explains, "This modern framework is ideal for realizing the fine-grained functionality we wanted, and it meshes well with the current trends in the academic community."

Like this, MEKO3 offers metadata handling with a high degree of flexibility, making the service easy to browse, search, and customize for every kind of user—general readers, registered users, and repository administrators.

Continuous evolution through both development and operation

Since migration to the next-generation JAIRO Cloud built on WEKO3 will affect over 600 user institutions, the migration will be implemented in phases, beginning in March 2021. A mountain of work still needs to be done, such as preparing migration documents, coordinating with integrated services, data conversion and checking by user institutions, and switching services. The essential thing is keeping the development and operations teams on the same page and working closely with the JPCOAR people who handle user support.

HAYASHI, Yutaka comments, "Development and operations cannot be completely separated. What is important is to communicate with each other carefully and continue to make small improvements from both perspectives." For both sides, the goal is to help shape an academic community that makes greater use of research findings. WEKO3/JAIRO Cloud will undoubtedly be a cornerstone of this effort.

Professor HAYASHI comments, "We are also working on agile development, but many things are not working out at the stage of trial and error. Nevertheless, we will keep tackling this challenge with the aim of accomplishing both rapid R&D and stable service operation." The trial-and-error process of supporting Japan's institutional repository service goes on.

Interview/Report by GOMI, Akiko   Photography by SATO, Yusuke

CiNii Research's new role as a knowledge search platform

Paving the way for open science, from citations to relations

NII's academic search platform "CiNii" consists of three services, "CiNii Articles" for searching for articles in Japan, "CiNii Books" for searching for books and journals kept in university libraries all over Japan, and "CiNii Dissertations" for searching for doctoral dissertations in Japan. Now, a new search platform called "CiNii Research" is under development. It will integrate these services and allow users to search for research data associated with books, articles, and dissertations. In November 2020, a preview version of the new integrated service was released under the name "CiNii Research Preview" (hereinafter "preview version").

ONAMI, Jun-ichi
Project Associate Professor, Research Center for Open Science and Data Platform, NII

SAKAGUCHI, Koji
Head, Academic Content Team, Scholarly and Academic Information Division, Cyber Science Infrastructure Development Department, NII

Combining three services to enable across-the-board search

The new CiNii Research differs from CiNii in two basic ways. Firstly, the three separate search services of CiNii are now integrated, and secondly, it is now possible to search for research data and research project information. (See Fig.) Infrastructure-wise, the CiNii Research is new, using almost none of the technology of CiNii. The search platform is built on a public cloud and boasts an excellent modern and high-performance architecture, with API links to external databases and overseas academic institutions, and search functionality built on the open-source "Elasticsearch" search engine.

Professor ONAMI, Jun-ichi, who leads the project development team, explains that the major revision was prompted by a rapid shift in the science world from citations to relations, driven by the need to trace the relations between articles, research data, and other relevant materials, and to increase opportunities for reusing research data in other studies. In recent years, as the concept of open science has penetrated the academic world, there has been a strong trend toward openly publishing research data. Furthermore, since the existing CiNii is built on specialized frameworks for each kind of resource, there have been growing demands for an integrated interface that enables cross-sectional searching. As a response to these changes and demands of the times, NII began the development of CiNii Research in 2017, working steadily toward an official release in April 2021.

Related research data and projects are now searchable

For enhanced functionality, CiNii Research features an original graph database. In a graph database, the entities of individual information items (e.g., articles, research data, bibliographies) are represented as nodes (or vertices) on a graph, while the relations between entities are expressed using edges (lines). Since there is no limit on the number or types of entities connected to an entity, adding data is easy, and searching for relations is faster and deeper. SAKAGUCHI, Koji, who has been involved in the project from the outset, has high expectations of the new service, saying, "Previously you ran a search for one entity and that was it. Now, if we can provide the ability to run successive searches across the board for other related entities, we can significantly accelerate the move to open science."
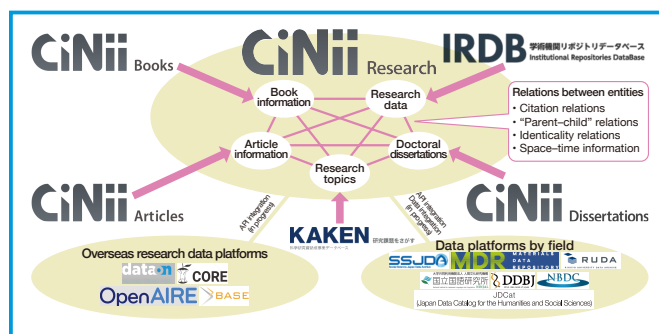

Fig: Scope of application of CiNii Research

Professor ONAMI explains that since each of the existing services uses a different schema (structure) for storing metadata, the developers had to grapple with the issue of standardization, to figure out how to store metadata under the new platform." In the end, for reasons related to ease of service integration, they opted to employ the "JPCOAR Schema" formulated by JPCOAR, which jointly operates the public "WEKO3/JAIRO Cloud" platform.

The preview version of CiNii Research that has been publicly released has been very well received. Typical comments about the service include, "This is the kind of search platform I was looking for"; "Searching is faster"; and "It's so convenient to be able to use a single search platform." The developers will be improving and updating the service on a continuous basis, taking on board user feedback.

Mr. SAKAGUCHI adds, "Although CiNii Research is a newly developed service, it retains the easy search and display functions of its predecessor. Implementing search functionality for Japanese is considerably more difficult than for English, so we had to work on optimizing this in CiNii." Thus, CiNii's mission of enabling users to search an abundance of academic information in Japanese in an easy-to-view format, has been carried over to CiNii Research.

Professor ONAMI describes CiNii Research as "a software extension of the modern librarian." As the heir of Japan's academic assets, CiNii Research is flowing with the tide of change, to make a fresh start as a search platform for Japan's collective knowledge.

Interview/Report by GOMI, Akiko　Photography by SATO, Yusuke

First training course in Japan for research data management

From system development to training support personnel

In June 2021, the GakuNin LMS (Learning Management System) for the GakuNin Academic Access Management Federation, which NII is building with universities nationwide, will officially begin offering training courses for institutions starting to manage research data. We asked two people who worked on developing the course with Japan Consortium for Open Access Repository (JPCOAR) about the state of research data management in Japan and the key points for connecting the training course to practice.

FURUKAWA, Masako
Assistant Professor, Information and Society Research Division; Research Center for Open Science and Research Data Platform, NII

MINAMIYAMA, Yasuyuki
Project Technical Specialist, Research Center for Open Science and Research Data Platform, NII

Knowing the importance of research data management and building a system suited to one's own institution

The work of developing the educational materials for the training course was started around 2015 by JPCOAR, a consortium of university libraries and other organizations that aims at publishing and distributing research data through institutional repositories, for the purpose of training personnel in research data management. Although the course contents reference research data management case studies from Europe and the U.S., they have been adapted to the circumstances of Japanese universities and research institutes. NII has collaborated with JPCOAR to provide teaching materials on a trial basis, through MOOCs and GakuNin LMS, a learning management system. MINAMIYAMA, Yasuyuki, a project technical specialist involved in developing the teaching materials, explains, "Libraries are typically expected to play a leading role in managing research data at universities, making use of their accumulated experience in organizing information and publishing papers ahead of data publication. However, since the standing and state of libraries vary between countries and institutions, it is essential to build a practical research data management system tailored to the specific situation. Library staff in Japan has never been directly involved in research data management, so the university's URA and IT center need to collaborate. In creating these materials, we have kept this point firmly in mind.

The training course starts with recognizing the need for research data management. Professor FURUKAWA, Masako, who has worked on everything from developing teaching materials to building the LMS, explains, "Up to now, researchers have managed their research data on their own, but if we can get universities and research institutes to provide them research data management support continuously on an organizational basis, we can promote open science and encourage more cross-disciplinary use of research data. And this will lead to new scientific knowledge. Our challenge is to get people to understand that research data management is a fundamental requirement for this."

A look at all the course contents offers an overview of the support (services) required during the research data lifecycle (generation → processing → analysis → storage → publication → reuse). (See Fig.) The course also makes clear the policies that need to be followed when managing research data at one's institution, the departments that need to be asked to collaborate, and the personnel and skills that need to be cultivated through training. With this basic knowledge, institutions can each set up and implement their own research data management system.

Continual development for better course delivery

Counting MOOCs, the training course has already been taken by thousands of people. Many of these reported that the course enabled them to understand research data management in a broad and systematic way. However, given that each institution requires different educational contents and because people want to learn as efficiently and quickly as possible, it was decided to offer the course in the form of "micro content" pieces of a few minutes each. This allows institutions to customize the contents by freely combining the pieces to suit their needs. Professor FURUKAWA stresses, "With this LMS, we can offer an environment that makes learning easy and course materials can be easily revised. On top of that, the ability to analyze detailed learning history data makes it understand what learners find difficult." Like this, the LMS offers a system for continuously providing up-to-date and highly effective training courses.

Looking ahead, the plan is to provide educational materials for researchers in the form of training courses. Mr. MINAMIYAMA says, "We want to help establish research data management so that it is seen not as an obligation for publishing papers or receiving grants, but rather as an activity that contributes to the convenience and growth of research." Professor FURUKAWA adds, "It is advisable for everyone involved in research, including students, to learn research data management as an element of research literacy." While the development of educational materials will continue in response to the demands of society, it will be interesting to see how much the training course, which officially opens to the public in June 2021, can boost research data management at Japanese academic institutions.

"Research Data Management in the Age of Open Science"
- Week 1: Behind the growing importance of RDM and the value of RDM
- Week 2: Storing, sharing, and documenting research data
- Week 3: Metadata and legal and ethical issues
- Week 4: Policy and service design for research data
The course videos can be viewed on the NII website: https://www.nii.ac.jp/service/jmooc/rdm/
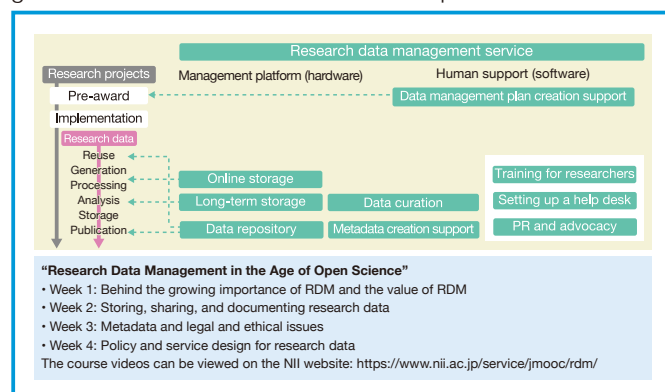
Fig: Overview of research data management service (top) and contents of training course "Research Data Management in the Age of Open Science" (bottom). Training course site: https://lms.nii.ac.jp/

Interview/Report by IKEDA, Akiko   Photography by Furusue, Takuya

The dawn of inter-university networks

ADACHI, Jun

Vice Director-General, NII

In the process of putting together "20 Years of the National Institute of Informatics," it occurred to me that the greatest contribution NII has made to universities has been SINET. In connection to this, I would like to write a few words on the early days of the inter-university network.

By the 1970s, the idea that could be considered the germ of our academic information infrastructure of today was already present. To make the idea a reality, though, large-scale joint research projects funded by grants-in-aid for scientific research were implemented in succession. The special research project "Higher-order Processing of Wide-area Mass Information" (from 1973), led by Professor SHIMAUCHI, Takehiko, director of The University of Tokyo's Computer Center, who is the first-generation "Dr. Ochanomizu" of this project, followed by Professor KITAGAWA, Toshio and others, gave rise to the concept of "academic information system." Then, with the special research project "Processes for Shaping Information Systems and Organizing Academic Information" from 1976, the leading roles were played by Professor SHIMAUCHI, followed by Professor. INOSE, Hiroshi, director of the University of Tokyo's Computer Center (and later director general of NII). It was at this point that the N-1 protocol was developed as a method for connecting computers to each other, centered on the work of Professor ASANO, Shoichiro.

In 1979, when this system was finally operational, Professor TAKANO, Kiyoshi of the University of Tokyo's Computer Center led the management and development of the system. As a graduate student, I helped out a little with connection experiments. Rather than ARPANET's TCP/IP, the communication method used X.25, a packet switching protocol designed for the lower three network layers, defined as a standard by CCITT. However, even though we connected up the large mainframe computer centers of seven universities, we were faced with the headache of how to cover the communication charges. Nippon Telegraph and Telephone Public Corporation's DDX packet switching service, which launched in 1980, cost ¥1 per packet. As a tentative fee plan, we decided to split the cost between the large mainframe computing centers and users, but there were no signs that usage was increasing.

As late as 1984, when we connected university libraries and started running the book catalog system, communication fees were still a thorny issue. The following year, the very energetic NISHIO, Masahiro was appointed to head the Science Information Division of the Ministry of Education, Science, and Culture. He proposed the bold idea of establishing an academic information center as an inter-university institution. Together with some others, I was pushed along by Mr. NISHIO's momentum to start writing requests for cost estimates for the creation of the center. During the discussions, Professor HASHIZUME, Hiromichi, who was then working as my assistant, hit on the idea of buying up all the high-speed optical fiber digital lines and X.25 switching equipment that NTT had started offering. The idea worked out splendidly. These are the roots of SINET, a self-managed network without any worries about usage-based charges.

While in the U.S. in 1986, I happened to meet former Bell Laboratories professor John Pierce. I boasted that we had started building a dedicated network in Japan using X.25 instead of T1. He quickly shot back, "How do you manage to get any speed with such a complicated protocol?" I was a little disappointed when I later learned that the network speed barely reached 256 kbps. Our academic information network has seen many twists and turns since then, such as connecting to CSNET and BITNET for international e-mail, providing lines to JUNET, and the rise and fall of ATM. Happily, though, since the turn of the 21st century, when NII took over, SINET has established itself as a solid network offering a high level of speed and reliability.

When I think back and realize that most of the communications terms used in this article are now long forgotten, I feel a little sad.

Future Schedule

June 14-19: Japan Open Science Summit 2021 = Online. Details at https://joss.rcos.nii.ac.jp/

June 18-19: National Institute of Informatics Open House 2021 (presentation of research findings, open to public)

Late June to early July (dates to be decided): National Institute of Informatics Academic Information Infrastructure Open Forum 2021

Details of both Open House and Open Forum will be announced on the NII website as soon as they are decided. https://www.nii.ac.jp/

Notes on cover illustration

Each of the robots is engaged in research, research data management, publication, and reuse. This open science cycle is expected to enable the acquisition of greater scientific knowledge and open up new research fields. The NII research data platform "NII RDC" strongly supports these activities.

Search "NII Today"!

Bit
(NII Character)

https://www.nii.ac.jp/en/about/publications/today/