National Institute of Informatics News ISSN 1883-1966 (Print) ISSN 1884-0817 (Online)

NII Today

National Institute of Informatics News

104
Jan. 2025

[Feature]

Research Data Ecosystem

Connecting and Circulating Research Data Across Disciplines

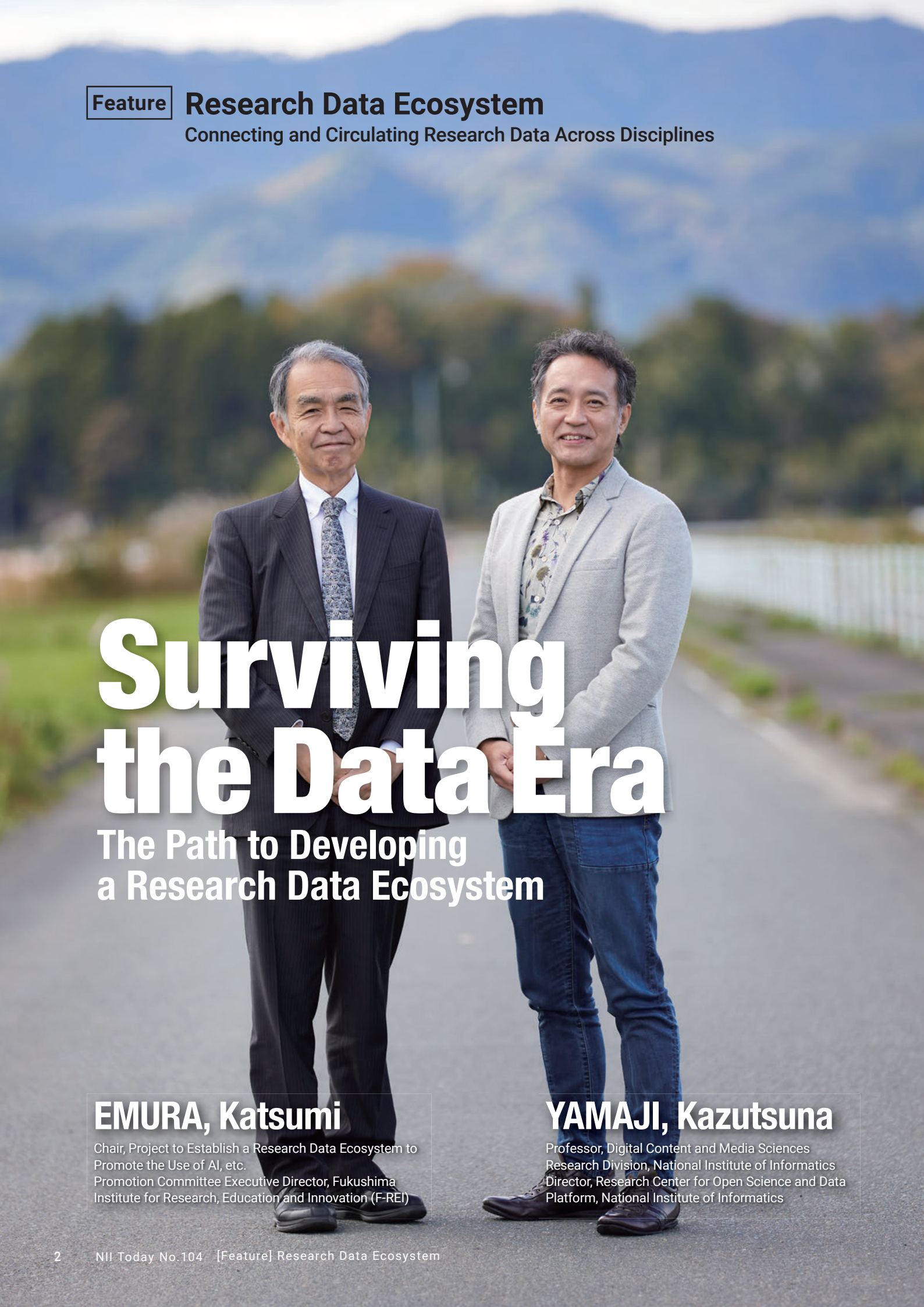
This English translated version NII Today corresponds to No.104 of the Japanese edition

www.nii.ac.jp/today

Feature

Research Data Ecosystem

Connecting and Circulating Research Data Across Disciplines



Surviving the Data Era

The Path to Developing a Research Data Ecosystem

EMURA, Katsumi

Chair, Project to Establish a Research Data Ecosystem to Promote the Use of AI, etc.
Promotion Committee Executive Director, Fukushima Institute for Research, Education and Innovation (F-REI)

YAMAJI, Kazutsuna

Professor, Digital Content and Media Sciences Research Division, National Institute of Informatics
Director, Research Center for Open Science and Data Platform, National Institute of Informatics

Research that aims to foster innovation by leveraging vast amounts of data, such as data-driven research and generative AI (artificial intelligence), is trending worldwide. The "open science" approach, which involves widely sharing not only research papers but also research data online, is also spreading internationally. The Project to Establish a Research Data Ecosystem to Promote the Use of AI, etc. aims to provide a common platform for enabling researchers all across Japan, regardless of their field, to engage in data management, publication, utilization, and application. We spoke with Project Promotion Committee Chair EMURA, Katsumi and Management Team Leader YAMAJI, Kazutsuna of the Center for Research Data Ecosystem Development, the project's core organization, to discuss the project's progress status and similar.

Sharing a Recognition Among Universities and Expanding the Community

Interviewer

TAKI, Junichi

Science Journalist
(Former editorial writer of the Nikkei)

—Roughly two years have passed since the launch of the Research Data Ecosystem Development Project in fiscal 2022. We are now approaching the halfway point toward the goal at the end of fiscal 2026. How do you rate the past two years?

YAMAJI: The National Institute of Informatics (NII) has been working on the development of the NII Research Data Cloud (NII RDC) in its aim to encourage the utilization and application of research data. The NII RDC is a comprehensive usage platform comprised of three platforms for managing, publishing, and searching for data. The Research Data Ecosystem Development Project involves the NII undertaking research and development to further expand and upgrade the functions of the NII RDC from seven perspectives, including utilization, reliability, and aggregation (Fig. 1). At the same time, the four institutions RIKEN, the University of Tokyo, Nagoya University, and Osaka University are promoting the development of a so-called ecosystem, that is, the necessary environment for research institutions to utilize research data platforms. This includes establishing rules and guidelines for data utilization and publication, as well as human resources development.

Collaboration between universities, and the formation of regional consortia

I feel that, now two years have passed, a more tangible approach is required for platform development. For instance, universities are of the opinion that they "want to have this kind of data policy, and for that, if we do not do X, we cannot fully realize organizational data management, so we want to do X first," or that they "need to also develop human resources." We are gradually coming to see the needs of universities in the management

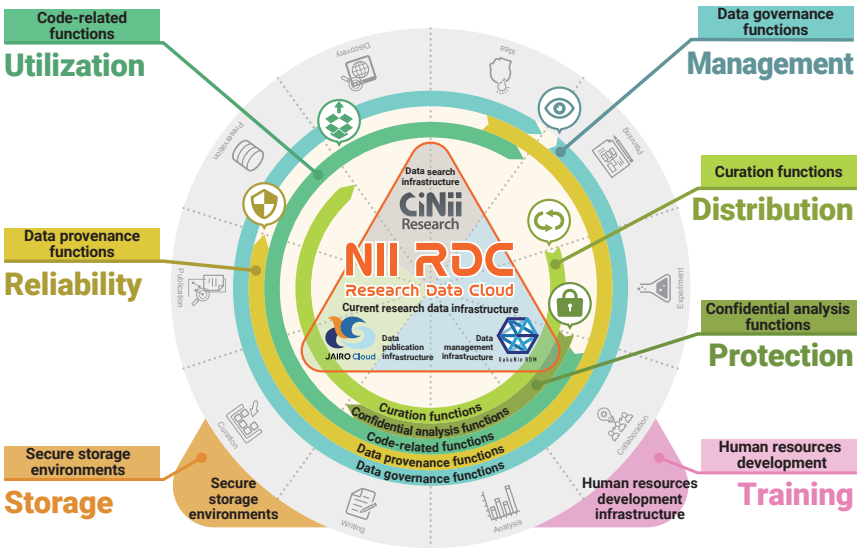


Figure 1 Conceptual diagram of expanded functions of the NII RDC (Research Data Cloud)
As one of the Research Data Ecosystem Construction Project's core institutions and the leading institution of the Research Data Infrastructure Enhancement Team, the NII aims to implement the functions of the research data platform by expanding the NII RDC from seven perspectives.

and utilization of research data. Toward its realization, efforts are being made toward collaboration between universities and the formation of regional consortia. Starting with the establishment of a consortium in the Tokai region led by Nagoya University in fiscal 2023, attempts are being made to create several such consortia in the Hokuriku, Chugoku-Shikoku, Kyushu, Hokkaido, and Tohoku regions. These efforts are not being initiated unilaterally on the project promotion side. It is also a movement emerging from universities with a similar awareness of the challenges.

We are making progress to upgrade the functions of the NII RDC to preempt these on-the-ground needs and adjust the project in line with the growing demands of universities, and the pieces are starting to perfectly fit into place.

The need for a shared data platform that can be used by researchers in Japan

—Why is there a need for the Research Data Ecosystem Development Project right now? Could you once again explain the background behind the project's inception?

EMURA: While our present time has been referred to as "the era of data-driven science" before, we have now fully arrived in the data era. Above all, the emergence and rapid proliferation of generative AI in recent years have led to a significant rise in the value of data.

Even before the arrival of generative AI, people were already talking about this being the era of open science and open data. The makeup of the scientific community is shifting from an era in which it was fine as long as researchers properly published papers, to encouraging open sharing of the underlying data supporting those papers.

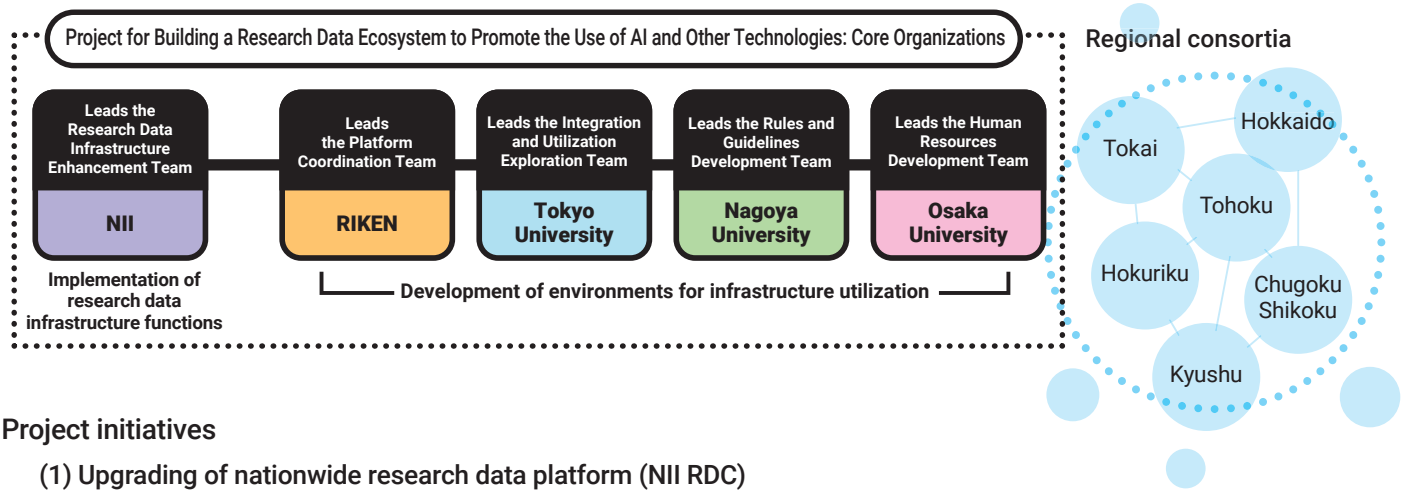
In July 2024, a joint statement was issued at the G7 Science and Technology Ministers' Meeting in Bologna, Italy advocating for the expansion of open science. Japan's Ministry of Education, Culture, Sports, Science and Technology has set out a policy to promote immediate and open access to academic papers toward facilitating the wider utilization of research results. Many researchers are now being called on to recognize the importance of open data and make more effective use of data.

In such an age, we cannot necessarily say that Japan is, frankly speaking, at the forefront globally from the perspective of data utilization. The first major aspect of the background behind the project is, considering that there is a strong societal and academic demand to properly establish a platform for enabling data utilization, to develop a common research data platform accessible to researchers nationwide, with NII at the core, in order that individual universities do not pursue the creation of various platforms in a more fragmented way.

The Research Data Ecosystem is a platform for data linkages Research Data Ecosystem

Another aspect is that we wanted to generate new value through data linkages. Up to now, it has been common for researchers to utilize data from within their own research area. Meanwhile, data linkages that connect diverse domains have led to the creation of new value. A platform is needed to make this a reality. That is what we call the Research Data Ecosystem. You could also say that this is the so-called ultimate goal of the project, so we intend to advance toward its realization.

—Creating new value through data utilization and application



Project initiatives

(1) Upgrading of nationwide research data platform (NII RDC)
(2) Development of environment for promoting usage of research data platform
(3) Strengthening of functions toward promoting open access

EMURA, Katsumi

Completed his master's at the University of Tokyo. Joined NEC Corporation as a researcher in optical communication. Served at NEC as Chief Technology Officer (CTO), senior advisor, and other roles before assuming his current roles. Has held numerous key roles, including member of the Information and Communications Council of the Ministry of Internal Affairs and Communications, Director of the AIP Network Laboratory at the Japan Science and Technology Agency (JST), and Vice-Chair of the Innovation and Technology Committee at Business at OECD (BIAC). Visiting Researcher at Bellcore, U.S., between 1987 and 1988. Doctor of Engineering, University of Tokyo.

is also important for enhancing Japan's research capacity, right?

EMURA: I think it's extremely important. It has been pointed out that the number of papers ranking among the top 1% and 10% published in Japan is decreasing, and Japan's research capacity is diminishing. I think that the core issue lies in the weakness of interdisciplinary research. According to analysis by the Ministry of Education, Culture, Sports, Science and Technology's National Institute of Science and Technology Policy (NISTEP), the number of academic fields gaining attention globally is increasing. Analysis from 2002 shows that there were about 600 research fields garnering international attention at that time, while analysis from 2020 shows that this number is over 900. The growth is particularly notable in interdisciplinary and transdisciplinary areas; that is, those that cross traditional academic fields such as physics and biology.

In countries including the United States and China, more and more papers are being published in these new areas and research is progressing, but Japan seems to be lagging behind. While Japan exhibits its strength in traditional fields, we can see a trend of little growth in the number of high-profile papers in interdisciplinary and transdisciplinary areas. I believe that this may be the root cause behind the overall decline in the number of top-tier papers from Japan.

In other words, Japan has struggled to create new academic disciplines. That is what makes this platform so critically important. If you were to say that having a platform will lead to the materialization of interdisciplinary fields, it's of course not that simple. Having said that, I have no doubt that the platform will enable a variety of people to come together and connect there.

Building communities through use cases is the starting point

—With different research fields, the nature and volume of target data will also be different. I hear that even establishing a single data policy can be challenging due to being unable to reach consensus within the university. Is human resources training also a major issue?

YAMAJI: As Committee Chair Mr. EMURA pointed out, creating interdisciplinary research areas and generating new value through data linkages can be said to be the ultimate goal of this project. We have been developing a platform aimed at the utilization and application of data as a precursor to this. The next thing that is required is to develop human resources fully versed in using the platform. Universities will likely need personnel to take on a coordinator-type role of introducing tools for managing, utilizing, and applying data to researchers. Overseas, personnel specializing in such roles are already boosting interdisciplinary research and

supporting data-driven research.
As part of the project, we put out a call for use case proposals, with 33 initiatives launched so far (as of January 2025). Communities using the same platform—in other words, "eating food from the same pot"—are forming, and as we support these communities, we expect to see more outcomes in future, such as advances in interdisciplinary joint research.

EMURA: Rather than seeing this as a challenge, we should recognize it as the path forward becoming clearer. A diverse range of researchers beginning to engage in activities using the same platform is the starting point. First, we need to build up communities. I believe that our perspective will truly shift once we reach the stage where it is possible for community members to propose that linking different datasets could generate value.

Creating interdisciplinary research areas through data linking

EMURA: One of the roles of the project promotion committee is to evaluate the outcomes of the project, but the committee has also expressed its commitment to accompanying the project from the standpoint of promotion. Committee members come from diverse backgrounds, including humanities and social sciences fields such as law. As a result, feedback such as "please make it so that we can use it as well" is being raised. We need to ensure that many people understand the value of the research data ecosystem. I believe that communicating the significance and importance of generating interdisciplinary research areas through data linkages, and thereby creating new value, is a really important aspect of this project.

YAMAJI: To add to that, whether it is open science or data governance, research institutions tend to see these as "obligations." For researchers, there is an aspect of "forced work." However, those who have offered to take part in the project's call for use case creation do so not out of obligation; rather, they are trying to incorporate the platform in order to promote research.
While the National Science Foundation (NSF), the National Institutes of Health (NIH), or individual universities in the United States work on developing a data platform, there is also ongoing discussion concerning the need for increased investment in research infrastructure at the federal level. Similarly, the European Union (EU) has made huge investments into the European Open Science Cloud (EOSC). That said, these are not platforms that can cover an entire country like the NII RDC. Japan is thus a pioneer in some respects. Our role is to continue taking a leading position in the world by evolving the project into a service for promoting research, rather than mere compliance.

—While the project has a set ending, the system itself will continue on. Is there anything the project promotion committee should accomplish in anticipation of the next 10 or 20 years?

EMURA: From a long-term perspective, the sustainability of the research data ecosystem is a concern. The question is in what way the system will become independent. This is not something that the NII can resolve alone. I think it would be better to have participating universities, and in future, companies, get involved, but ultimately, the system's beneficiaries should be the ones responsible for maintaining it. If the system brings about new value, everyone will naturally want to participate. How this process is designed will be a critical factor. If you were to ask whether we could fully design this within the next two years, I honestly think that it won't be an easy task. Having said that, I do believe that the project promotion committee should also put forward suggestions as to in what way we can ensure the sustainability of the system.

A Word from the Interviewer

The Government of Japan published its slogan of "Data Free Flow with Trust (DFFT)" in 2019. Looking back at the subsequent response to the COVID-19 pandemic, I felt that the road to achieving this vision would be a long one. You could also say that data management and utilization at research institutions are now urgent priorities. The emergence of communities that share this recognition is encouraging. The challenge now is how to expand, deepen, and sustain these efforts.

TAKI, Junichi
Science Journalist

Joined Nikkei Inc. after graduating from Waseda University's School of Political Science and Economics. Has reported on topics such as science and technology research and development sites, science and technology policymaking processes, climate change, and energy policies since the mid-1980s. After leaving Nikkei Inc., served as Executive Secretary of the Japanese Association of Science and Technology Journalists.

Photos: MORI, Takahiro

How Much Progress Has Been Made in Developing a Research Data Ecosystem?

NAKANO, Keiichi
Program Manager
Center for Research Data Ecosystem Development
National Institute of Informatics

The Project to Establish a Research Data Ecosystem to Promote the Use of AI, etc. is an initiative led by the National Institute of Informatics (NII) in partnership with four research institutions/universities. The project, which was launched in 2022 and will conclude in 2026, is now a little past its midpoint as of January 2025. Its aim is to enhance the existing research data platform (NII RDC) while evolving it into an ecosystem capable of contributing to creating innovation by promoting the utilization and application of research data.

Its greatest significance lies in dramatically improving the transparency and reproducibility of academic research by promoting data sharing, utilization, and application. In an era where data analysis utilizing AI and machine learning is becoming more sophisticated, thus advancing data-driven scientific research, proper data management, utilization, and application are essential. It is crucial that data can be found easily by anyone (findable), can be accessed as necessary (accessible), can be linked with other systems and data (interoperable), and is offered in an easily reusable form (reusable). Further, it is also vital to ensure the protection of confidential information. Toward the creation of such a society, project members are engaged in upgrading the NII RDC, platform collaboration, creating use cases, developing rules and guidelines, and human resources development.

Platform collaboration involves examining technical or operational innovations necessary for organically linking the NII RDC with platform systems for other fields, such as those seen in disciplines where the open science approach is particularly prevalent, and with existing systems held by various institutions.

As for use case creation, a total of 33 topics (as of January 2025) from diverse research fields have been adopted, and are beginning to accumulate as practical examples of data utilization. These use cases are expected to lead to the discovery of newer research approaches, thereby creating an ecosystem where such research developments enhance and accelerate subsequent research.

Many institutions are highly interested in developing rules and guidelines. This is because in the 6th Science, Technology, and Innovation Basic Plan, all national universities, national research and development agencies, and other organizations with institutional repositories have been urged to formulate data policies by the end of 2025, and the deadline is fast approaching. If policies do not align with the specific circumstances of each institution, they will be of little use, but creating them from scratch is a daunting task. Despite this, leading institutions are publishing data policies tailored to their own needs, and the knowledge gained through these efforts is starting to be accumulated and shared. The groundwork for rolling out research data management nationwide is steadily being laid.

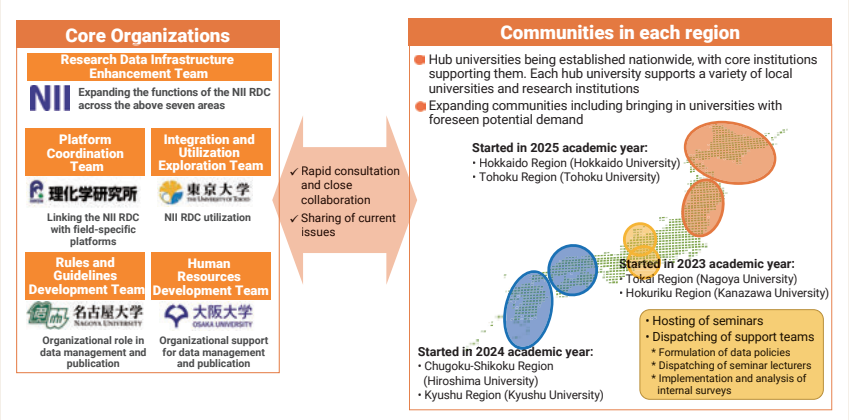
Human resources development initiatives involve developing programs to help individuals acquire the skills and knowledge necessary for research data management in an effective manner. If we develop personnel responsible for research data management alongside researchers, data literacy will improve, thus encouraging utilization of the ecosystem.

Moreover, in fiscal 2023, the Research Data Management Start-up Support Project was launched as part of the project. It offers initial support for regional research institutions and universities working on research data management. Unique use cases leveraging the strengths and resources of regional communities have already emerged, with data management practices tailored to each region's needs underway. As even more regional communities are established and form a national network, research institutions across the country will eventually evolve together, which will likely lead to the forming of a sustainable research data management platform.

The research data ecosystem is not just for researchers—it will also bring great value to society as a whole. By developing this ecosystem, tangible outcomes are expected across a diverse range of fields, such as solutions to regional issues, the development of industry, speedier disaster responses, and moreover, improvements to medical and welfare services. Our aim is to contribute to the realization of a richer and more sustainable future in which our entire society can enjoy the benefits of advanced research data utilization and application.

We hope to continue working with a variety of stakeholders in pursuit of the challenging yet crucial goal of developing a research data ecosystem so that we can contribute to the development of scientific research.

Structure and rollout status of the Research Data Management Start-up Support Project



Composition: MIWA, Yoshiko


Aiming to Transform Research Activities

Rules - Guidelines Development Team
Human Resources Development Team

The Project to Establish a Research Data Ecosystem to Promote the Use of AI, etc. of the Ministry of Education, Culture, Sports, Science and Technology, a project to promote the transformation of research activities through digital technology and data utilization—digital transformation of research (research DX)—involves the integrated promotion of the following: the formation and dissemination of usage cases, the acceleration of data sharing and utilization, and the effective use of digital infrastructure and similar for research. The project was launched in 2022 by the National Institute of Informatics (NII) alongside joint core institutions RIKEN, the University of Tokyo, Nagoya University, and Osaka University. In this context, we asked members of the Rules - Guidelines Development Team and Human Resources Development Team about their progress to date, challenges, and future outlook in their aim to support proper research data management and development of human resources for data management in order to accelerate data sharing and utilization at this time.

MATSUBARA, Shigeki, KAI, Naoto,
NAGAOKA, Chikako, HIRAKI, Toshiyuki


■Aiming to make research data management universal

——Could you provide an overview of the activities being conducted as part of the Research Data Ecosystem Development Project?

MATSUBARA: The Research Data Ecosystem Development Project is a collaborative effort between the NII and four other institutions. As the leading institution of the Rules-Guidelines Development Team, Nagoya University is working to develop and disseminate rules for handling research data at universities in Japan, aiming to create a unified national framework under which research data is handled.

At present, the handling of research data at universities is often left to the discretion of individual research groups or researchers. This offers researchers a great deal of flexibility, which may seem ideal at first glance, however having to make decisions about how data is handled for every action is inefficient.

Establishing rules and guidelines in advance will remove uncertainty for researchers in handling research data, thus streamlining procedures, and eliminate ambiguity regarding the significance, management methods, and the like of each data set when sharing research data in groups. There are also many advantages for research assistants, such as helping clarify how they can assist researchers.

Unifying data handling practices, storage locations, and the like, which varied across individual groups and researchers, will give universities a clearer understanding of research data governance; that is, what data they have and where it is stored. It will also make it easier to come up with data utilization policies, which I believe will be a major benefit for universities as well. As data from various research fields accumulates in one place, connections between people, between researchers, will strengthen through data utilization, which will surely enrich research activities. I believe this will, in turn, enhance the value of the data.

HIRAKI: I am collaborating with the Rules-Guidelines Development Team to develop functions that will support data governance. Proper research data management requires a clear understanding of the current state of data management. The data governance functions we are working on largely aim to realize the three pillars of 1. visualizing what we need to protect (policies and rules) in data management, 2. visualizing the current state of data management; and 3. helping to verify whether the current state of data management aligns with the policies and rules.

KAI: First, it is crucial to ensure that the importance of research data management is being widely and thoroughly disseminated. To this end, the Human Resources Development Team is working to establish a training environment giving researchers and research assistants a proper understanding of research data management. Specifically, we aim to first develop educational materials on data management and distribute them to research institutes nationwide through a learning management system (LMS). We are also working to develop occupation-specific learning curricula, along with building a learning analytics (LA) platform to provide effective learning environments tailored to individuals, aiming to train human resources capable of appropriately speeding up the cycle of the research data ecosystem.

NAGAOKA: We are utilizing GakuNin LMS, a learning management system provided by the NII, as a platform for distributing educational materials developed by Osaka University and standardized by the Japan Consortium for Open Access Repository (JPCOAR) nationwide. Through a range of platform functions built thus far, including providing learning content, learning analytics for making improvements, and a repository for reusing educational materials as a related service, the GakuNin LMS system will offer support related to human resources development.

Research data management (RDM) training can equip researchers with the knowledge and skills necessary for handling research data properly and ensuring it is in a reusable form. For this reason, I would like to position this training as an essential qualification for handling data, not just for researchers in certain fields, but for all researchers who handle research data. Therefore, our role is to provide a platform to support initiatives by Prof. KAI and Osaka University to create and disseminate educational materials.

■Current progress and challenges with project promotion

——Could you tell us about the current progress of your activities and any challenges you are facing?

MATSUBARA: I believe that promoting research data management at universities should involve a large number of universities cross-referencing with one another. In fact, when the Research Data Ecosystem Development Project was launched in the summer of 2022, fewer than 10 universities had established a research data policy, which can also be said to be the starting point for establishing rules and guidelines. As a national measure, the government has set a target for universities to establish a policy by the end of 2025, and as of the end of 2024, around 90 universities have done this, indicating that steady progress is being made.

In respect of another aspect, the publication of research data, universities nationwide are working to enhance their institutional repositories. The main focus was originally on publishing papers,

MATSUBARA, Shigeki

Professor, Information Technology Center, Nagoya University

but progress is now being made in creating metadata standards for publishing research data at repositories as well.

Meanwhile, we feel that this will be difficult to achieve in cases where practices have already become well established in research groups. These have likely been established as a result of optimization in various ways, making it difficult for researchers to abandon their current methods and shift to following university guidelines. I also think that many researchers and research assistants will find it a considerable burden to also publish research data. This is another challenging aspect.

The development and dissemination of the NII Research Data Cloud is expected to significantly contribute to eliminating these challenges.

HIRAKI: In respect to data governance functions, in the past two to three years in particular, we have been working to develop core functions; in other words, setting policies so that these can be properly determined by the system, and verifying the status of data management in accordance with the set policies.

We are finding it extremely challenging, however, to apply actual rules and guidelines as policies on the data governance function side. This is because, while content developed as rules and guidelines is understandable to humans, it is hard to read from the perspective of the system; in other words, its machine-readability is poor. There are times where we struggle with how best to enhance machine-readability and how we should build logic. I believe that we should work out these issues by discussing them with Prof. MATSUBARA.

NAGAOKA: With regard to human resources development, I feel that the creation of educational materials is progressing relatively smoothly. On the other hand, the biggest challenge from now on is likely to be how to disseminate research data management culture and education at each institution.

While I believe that knowledge about research data management and related educational materials will continue to increase in future, for example, even if the use of these resources were to be made mandatory, what level of study should be required of learners? Moreover, while I bring up the word 'mandatory,' it's not that we want learners, that is, researchers and research assistants, to feel that they are participating against their will. Rather, we want to make it so that learners can feel that there are benefits to learning and applying the acquired knowledge. While making these materials mandatory, we also want to explore such routes that will enable learners to actively engage with the material while having fun.

KAI: I believe that we have been able to make progress with creating and distributing educational materials faster than initially anticipated. I largely attribute this to how Osaka University and the NII have successfully worked as a team to share information in an exhaustive manner.

In particular, the fact that we began by setting the use of the NII's GakuNin LMS as a prerequisite for distributing educational materials has likely led to reducing the burden for target institutions.

NAGAOKA: The NII had been developing GakuNin LMS, a platform that enables users to share and access educational content, before the Research Data Ecosystem Development Project even began. This was done with the aim of reducing costs connected with developing and maintaining educational content at each institution, along with enabling the sharing of high-quality educational content created by certain educators and researchers on a national level. In a sense, the platform is a "seed" that has been cultivated for a long time, and I believe that it has been perfectly matched with this project.

■Vision for nationwide and horizontal rollout

——In light of these challenges, how is the outlook for the future?

MATSUBARA: The Rules-Guidelines Development Team plans to develop research data guidelines within fiscal 2024 and start distributing them nationwide from fiscal 2025. At the same time, as Mr. HIRAKI mentioned earlier, the question is how we should incorporate the developed guidelines as a function in the NII Research Data Cloud. We hope to make progress with this in collaboration with everyone at NII.

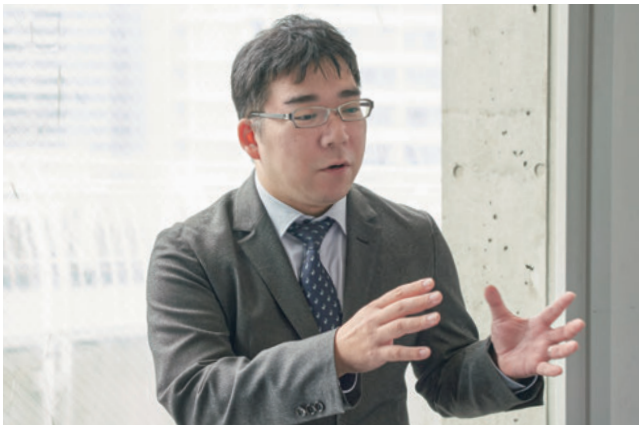
HIRAKI: We intend to promote

HIRAKI, Toshiyuki

Project Assistant Professor
Research Center for Open Science and Data Platform
National Institute of Informatics

KAI, Naoto

Associate Professor,
D3 Center,
Osaka University



the development of frameworks for making use of the formulated rules, guidelines, and policies in research data management. Are those policies functioning as a system within the data governance functions we are currently developing? Moreover, can researchers be managed under those policies? I believe it essential to ensure that these are in a form that can be fully verified. To achieve them, we hope to also make progress in examining improvements to machine-readability of policies and so on.

KAI: As I mentioned earlier, our target in respect of human resources development is the nationwide and horizontal rollout of educational frameworks, such as the materials currently under development. In reality, the type of human resources development needed to effectively maintain the research data ecosystem will vary by field and domain. Having said that, in terms of research institutions taking the initial step of first disseminating a guiding philosophy, I think that the things that need to be done are roughly the same for all institutions.

In future, to achieve significant results as efficiently as possible, I believe it necessary to establish an environment for developing human resources that can be adopted universally by research institutions and universities nationwide. As for relationships with other teams, we are currently collaborating with Nagoya University.

Researchers today are obligated not only to learn about data management, but also research ethics, information security, and various other topics. In such a context, in order to allow them to set aside more time for their original research activities, too, it is essential that they can learn with as little burden as possible. As a way of making learning less burdensome, we will not, for example, require everyone to carefully view all educational materials currently under development, so I think it will be essential to devise efforts including creating curricula optimized for individuals in line with respective needs, such as job roles, and allowing learners to focus only on the parts necessary for them. We aim to achieve this by incorporating the required rules and guidelines at each phase of the research data lifecycle in which each researcher and research assistant is deeply involved. We are hoping to coordinate with Nagoya University in this regard.

NAGAOKA: As for the outlook for human resources development, we are also considering using Open Badges, which are starting to be adopted at a wide range of universities, as a system for visualizing where human resources are located and which knowledge and skills they possess.

In addition, as Prof. KAI mentioned, we are also promoting curricula optimized for each individual. Circumstances vary for each team and each institution. We intend, wherever possible, to provide unified and accessible educational materials as a packaged offering; while this is fine for introductory resources, we also hope to work on creating a system that will allow each institution to customize materials to fit their respective needs.

MATSUBARA: Rule and guideline development and human resources development are the two crucial elements for research data management at universities. It is essential that we do not just establish rules, but also develop the human resources needed to appropriately instill and enforce these rules, along with creating related educational materials. We will continue to strengthen relations between the two teams in promoting our activities.

NAGAOKA, Chikako

Project Assistant Professor
Research Center for Open Science and Data Platform
National Institute of Informatics

Interview and composition: KAWABATA, Hideki/Photos: SUGISAKI, Kyoichi

Aiming to Pioneer the Use of the Research Data Cloud

ONO, Kanta
Professor, Graduate School of Engineering, Osaka University

KIMURA, Eizen
Professor, Graduate School of Medicine, Ehime University

TOKUCHI, Naoko
Professor, Division of Forest Ecosystem, Field Science Education and Research Center, Kyoto University

KIKUCHI, Nobuhiko
Associate Professor, Research Department, National Institute of Japanese Literature and Deputy Director, Center for Data-Driven Research on Premodern Japanese Texts

With the Project to Establish a Research Data Ecosystem to Promote the Use of AI, etc., proposals are being solicited for creating pioneering use cases in promoting the usage of a nationwide research data platform and data linkage.
The phrase "using a data platform" might sound simple, however many differences are emerging in terms of awareness of the challenges involved and approaches to utilization across different fields and research topics. Here, we talk with four researchers, each from different fields, whose project proposals were selected for the Use Case Creation Project (below, Use Case Project). We ask them about what influenced them to apply for the project, their progress, challenges, aspirations, and vision for the future of research data, and recommendations they have for the National Institute of Informatics (NII) based on their experiences participating in the project.

Awareness of challenges related to the Research Data Cloud

——Could you tell us the history of your participation in the current use case project and give an overview of your research?

ONO: Our proposal falls under the "experimental" category. The research we are conducting in the field of materials science can be broadly split into two types. One involves using large-scale facilities to acquire and analyze huge volumes of measurement data on synchrotron radiation and neutrons, and to discover something new from the data. The other is what we call automated and autonomous experiments. This involves tasking robots with carrying out experiments without any human intervention, which also enables us to continuously generate large volumes of data.
Currently, however, the focus tends to remain solely on acquiring large volumes of data. The loop of quickly analyzing these large volumes of data and applying these insights in subsequent experiments, which should have been valued from the beginning, has not been implemented as it should be. This had been a concern for me for some time.
At that time, I learned that the NII was just about to launch an initiative related to a data platform,

which led me to apply for this project. When applying, my proposal included building a platform for data sharing within the researcher community and linking this with the creation of new knowledge (Fig. 1).

KIMURA: My proposal set out the topic of examining a medical information analysis platform for realizing international real world data (RWD)*1 research (Fig. 2). Medical real world data is the general term for the wide range of healthcare data obtained in everyday clinical practice.

While working as a university professor with a specialism in medical informatics, I am also involved in operation and management of the electronic medical record (EMR) system at the school's affiliated hospital as head of the Department of Medical Informatics. At the same time, I receive requests from a large number of physicians who want to extract and use data from electronic medical records, but in reality, this is not a straightforward process. The data cannot be used immediately after extraction; rather, proper management is important, including the process of anonymizing personal information.

I had already been working on the development of a data environment platform for streamlining such management and usage, however, perhaps because my efforts were categorized under teaching or administration rather than research, they did not receive much public attention. It was under such circumstances that I put forward my application. The project has been accepted and is underway.

TOKUCHI: I belong to Kyoto University's Field Science Education and Research Center. The organization was created about 20 years ago during a university reorganization with the mission of "Connectivity of Hills, Humans and Oceans (CoHHO)," its newly established academic field. As the name suggests, this involves considering the environment from the forests to the seas as an interconnected whole in an effort to conserve national land and keep it healthy.

I originally specialized in forest ecology, but since then, I and other researchers at the center have broadened our purviews to carry out research beyond our original specialisms. In essence, we have been conducting what is now referred to as transdisciplinary research ever since then.

However, while natural science researchers like us tend to think of things in terms of quantitative value data, social science approaches can be more narrative based. How can we reconcile these different data "approaches"? This has always been a concern for me. When I brought up this issue with a professor in the field of informatics, they told me about this project. That was what prompted me to apply.

At that time, I just so happened to be conducting research on the topic of the forests in Hida City, Gifu Prefecture. This led me to propose the subject of "Can a platform be built that will enable us to consider the future of Hida City based on city data?" (Fig. 3). As for further ahead, our ultimate goal is to create a platform that brings together people from different fields to examine issues together. We are definitely not there yet, though.

KIKUCHI: I work at the National Institute of Japanese Literature (NIJL). The institute recently launched a 10-year Ministry of Education, Culture, Sports, Science and Technology Large-scale Academic Frontiers Project titled "Model Building in the Humanities through Data-Driven Problem Solving: Pioneering Next-Generation Humanities Research by Building and Utilizing Data Infrastructure," which will run between fiscal 2024 and fiscal 2033. As part of this project, the management of humanities research data and its publication and distribution as open data are pressing issues. I learned about the Use Case Creation Project through my supervisor. It seems like a great opportunity to gain experience in managing and publishing humanities research data.

Our proposal is titled "Research Resource Construction Project for Data-Driven Humanities Using Textual Data from Pre-Modern Japanese Texts." Using the Union

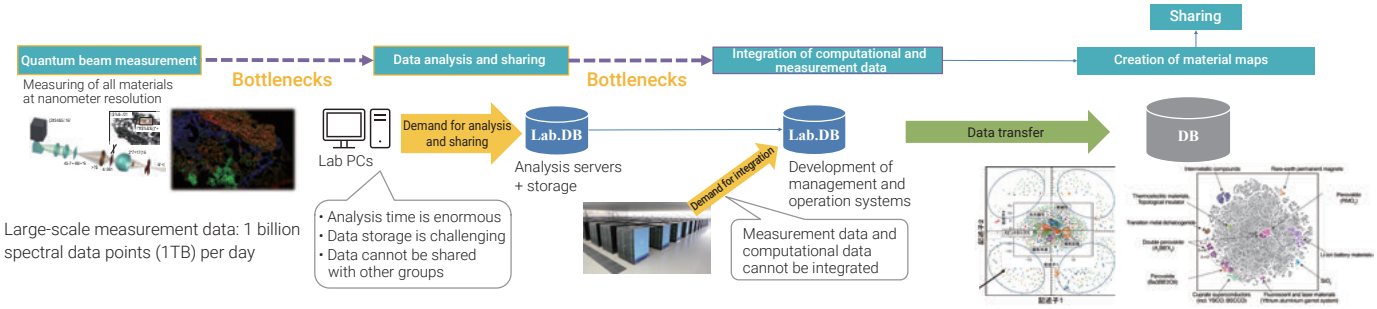


Figure 1 "Creation of knowledge through measurement, analysis, sharing, and publication: aiming to develop an ecosystem for the creation of knowledge through the measurement, analysis, sharing, and publication of large-volume measurement data using the NII RDC, GakuNin RDM, and mdx," from Professor ONO's poster presentation at the Research Data Ecosystem Symposium 2024.

Catalogue Database of Japanese Texts, which has been developed over the past decade, we are working to digitize distinctive classical texts in the TEI/XML format, guided by the research interests of participating team members. By sharing and publishing the insights and data gained through this trial-and-error process via the NII Research Data Cloud (NII RDC), we aim to transform classical texts into research resources (Fig. 4).

Varied levels of NII RDC usage

——Could you tell us about the requirements of the Use Case Project and how you are making use of NII RDC, including any benefits and challenges you've encountered?

KIMURA: For my research plan, within the NII RDC, I am currently using the management platform GakuNin RDM in particular. I find GakuNin RDM valuable especially in terms of ensuring the fairness and traceability of research.

In research, an extremely important aspect is whether or not its past processes have been transparent. GakuNin RDM provides a timestamp function and links with a highly trustworthy authentication platform. I think it is a highly useful system as system architecture that can ensure third-party credibility. From the perspective of patients, however, there may be a sense of resistance to the idea that storing personal information, like medical and research data, is really okay on such a widely used platform. Ultimately, this is a question of trust, but I believe it necessary to also conduct outreach activities to boost public confidence. I guess that is one of the challenges I can think of.

KIKUCHI: As part of our research, we use the NII RDC to store and share TEI/XML markup data with team members involved in joint research. We also use it as a place to consolidate those markup policies. Since the data we handle is often in small volumes, and we do not need to share data constantly, but only during research conferences, commercial storage or sharing via email is sufficient at present. In that sense, the benefits of using the NII RDC are not all that apparent, to be honest. Despite that, I believe that an advantage is definitely the peace of mind that comes from using the NII RDC. While there is no guarantee that conventional commercial cloud services will stick around permanently, the NII's system offers the reassurance of secure, long-term availability.

Having said that, this security, versus the drawbacks and areas for improvement, are simply opposite sides of the same coin. The ultimate sense of security comes from not being overly conscious of its presence or use, however using the NII RCD requires actions outside of our usual research practices, registering data using unfamiliar operations, and managing different data versions. Storing and sharing data used to be possible with a single click using commercial storage servers, but with the NII RCD, we are now tasked with managing metadata, and I can't help but feel that that makes things more complicated. I hope these hurdles are lowered.

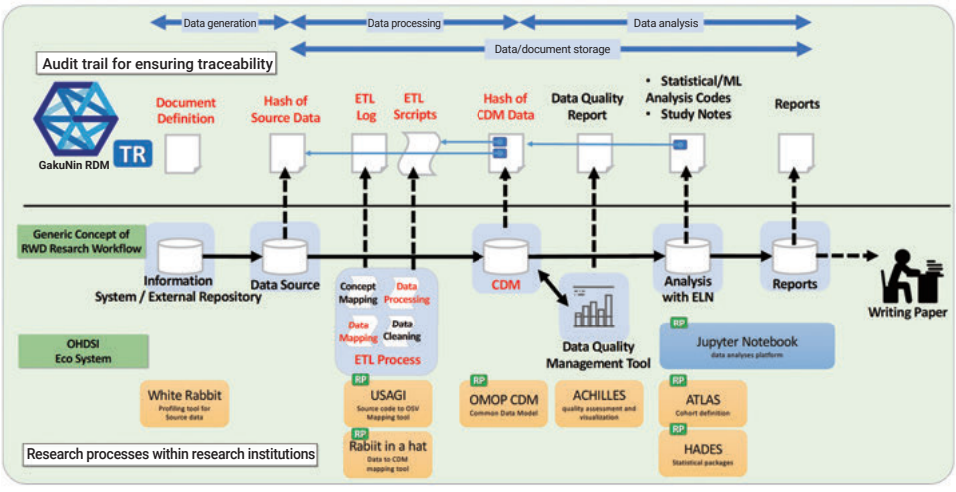
TOKUCHI: At the research center, we are only just getting started with using the NII RDC. This is because our project involves professors from various academic disciplines and institutions, and many of these institutions have not joined GakuNin. We had to start from the position of how to get them to join up, which was pretty difficult.

Finally, in counsel with faculty members, we managed to have everyone join as "affiliated researchers" of Kyoto University. It took a lot of time to get to that point, and we are not yet at the stage where we can say we are "utilizing" the NII RDC. As we move forward in future, I hope the platform becomes easier to join up for a wider range of people, if possible allowing individual researchers to apply directly.

ONO: At our university, we are only at the level of analyzing large volumes of data and sharing it using storage. We have not yet reached the point of managing all our research data on the NII RDC.

Has the NII RDC been designed for use as a place to store the vast volumes of data we handle? And, while I see no issues at all using it within our university, what should we do when we want to collaborate with external partners? For automated and autonomous experiments, we would like to carry out data accumulation and all

Figure 2 Investigating and categorizing an OHDSI*2 ecosystem, an emerging trend in RWD*1 research, to establish a foundation for defining metadata and exploring best practices for research using [GakuNin] RDM," from Professor KIMURA's poster presentation at the Research Data Ecosystem Symposium 2024.



other operations, including metadata-type data, without any human intervention at all, but how can we achieve this? There are still points that we are unclear on. This, too, is a challenge for the future, and I plan to go forward by looking into it further and consulting with members of the NII.

Toward future development

—What about your future ambitions and outlooks?

KIKUCHI: What we talk about in this project as a "research data ecosystem" is mainly targeting research data generated as part of future research activities.

From a humanities perspective, however, I would like to see consideration also given to distributing research findings that have already been published commercially as research data. Traditionally, the standard has been to publish research results as printed books, and this remains the case even now, but turning these results into printed books makes it more challenging to utilize the included content as research data.

In any case, researchers in the humanities are not yet accustomed to the concept of "research data management." I hope that the creators will do their best to develop the NII RDC system into one that is accessible for those working in humanities fields.

TOKUCHI: I don't have any particular grand outlook such as to "develop a research data ecosystem," but regarding solutions to the challenge in Hida City that we proposed as a project, I hope first of all to organically connect the data we already possess, and from this, realize a way to utilize the data in such a way that reveals insights into the nature of Hida City.

KIMURA: At this stage, I believe that the elements of data publication, data management, and researcher information management remain independent from each other and are not yet fully integrated within the NII RDC. My biggest wish is to see the steady integration of these elements. Moreover, in the medical field, there is often hesitation about publishing research data, which includes handling personal information. In respect of this, guidelines on pseudonymization have been published just recently, but it would be quite a burden for researchers to respond to each of these kinds of developments as individuals. If, as part of the service, there was a process by a third-party organization to assess whether data can be safely disclosed, it might further accelerate adoption.

ONO: As I mentioned at the beginning, in the field of automated and autonomous experiments, the aspect of data acquisition using robots and similar technologies is at the forefront, and overseas research is also focusing on this area. I think participating in the Research Data Ecosystem Development Project has made it clear that we need to go beyond building automated and autonomous hardware to reach the point of creating something that is truly new and valuable by working on how best to manage and utilize the data side. This is something lacking not only in Japan, but also globally, and it will continue growing in importance. What kind of data platform will be necessary to achieve this? I plan to work together with everyone at the NII in considering and building up such a platform.

Figure 3 "Areas covered in the present research: 1. gathering of materials that can contribute to community design; 2. development of a data platform; and 3. utilization and development of GakuNin RDM," from Professor TOKUCHI's poster presentation at the Research Data Ecosystem Symposium 2024.

In the present research project, the team will:
1. Generate text data from classical text image data and convert it into the TEI/XML format.
2. Through our research activities, aggregate and publish knowledge on creating classical text TEI/XML data.

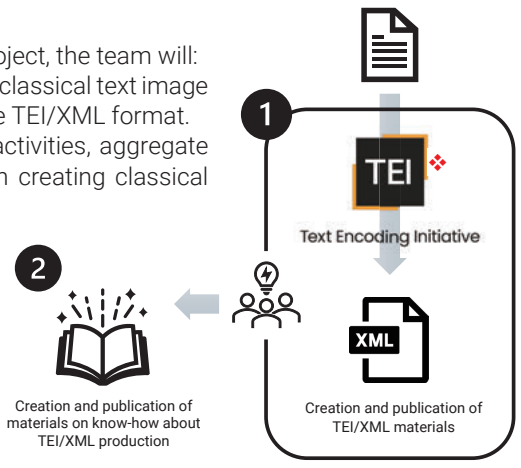


Figure 4 "Aiming to establish research resources for data-driven humanities by disseminating information not only to researchers of Japanese literature, but also to institutions holding classical texts both domestically and internationally, as well as to researcher communities in different fields," from Assistant Professor KIKUCHI's poster presentation at the Research Data Ecosystem Symposium 2024.

❖https://tei-c.org/about/logos/

*1: RWD (real world data): Anonymized medical big data that includes records of symptoms, prescription histories, and other information based on clinical practices.
*2: OHDSI (Observational Health Data Sciences and Informatics) An international, volunteer-driven open science community focused on medical big data analysis.

Policy as a Catalyst, Management as a Driving Force

Initiatives by Kanazawa University

In fiscal 2023, Kanazawa University was selected as a regional hub institution for the NII's Research Data Management Start-up Support Project, taking on the role of providing support to other universities. Starting from fiscal 2024, Kanazawa University will lead the Hokuriku Research Data Infrastructure Consortium alongside the University of Fukui, the Japan Advanced Institute of Science and Technology, and Kanazawa Medical University. In terms of the consortium's approach and framework for developing a pioneering research data ecosystem, even in Japan, and promoting research and development that leverages the unique characteristics of the Hokuriku region, how did this emerge, and what are the ultimate aims? We spoke to NAGAI, Keiji, Associate Professor at Kanazawa University.

NAGAI, Keiji
Associate Professor/
University Research Administrator,
Frontier Science and
Social Co-creation Initiative,
Kanazawa University

A pioneering spirit for data management

"Kanazawa University has long been highly conscious of the need for proper management of data generated through research activities. For instance, in 2020, when we were selected for the core facility program, which promotes research activities through the strategic sharing of research facilities among universities, we also recognized the need for proper data management, and therefore adopted and began trialing GakuNin RDM. I believe Kanazawa University was one of only two institutions working on the coordination of core facilities from the beginning," explains NAGAI, Keiji, Associate Professor at Kanazawa University's Frontier Science and Social Co-creation Initiative. As a chemist, Prof. NAGAI has conducted research in diverse fields, including new materials and laser fusion. He has experienced the difficulties of communicating across different disciplines and standpoints, the joy of overcoming those difficulties to collaborate successfully, and the power of cultural fusion in producing results. After achieving numerous accomplishments in the field of light sources for semiconductor lithography, Prof. NAGAI assumed his current position at Kanazawa University in 2021. He now works as a university research administrator (URA) to support and promote research activities. He joined the Ministry of Education, Culture, Sports, Science and Technology's Core Facility Construction Support Program from the second year of the five-year project.

"In order to appropriately manage and utilize data, we adopted the research data management platform GakuNin RDM provided by the NII and began conducting trials, and in 2022, we formulated the Kanazawa University Academic Data Management Policy. I believe that this is what led to the NII's Research Data Management Start-up Support Project (below, Start-up Project), which was launched in 2023" (NAGAI).

Kanazawa University and the NII have been engaging in joint research for more than a decade, which has yielded countless results. One such result is the digital distributed shared system for use among academic institutions. While universities across Japan were developing institutional repositories, Kanazawa University was already taking the next step.

Aiming to be an example for Hokuriku, and the rest of Japan

There was context behind Kanazawa University taking an important step in developing and promoting its research data ecosystem.

"There was momentum toward publishing data first of all at Kanazawa University's Innovation Airport. Moreover, the Core Facility Project also showed a direction toward promoting data science and the utilization and application of data. The Start-up Project was launched once these two significant developments were put in place (Fig. 1). I believe that was a crucial point. Our goal is to become a model for Hokuriku and, if possible, for all of Japan" (NAGAI).

Kanazawa University has a history of initiatives relating to ICT systems underpinning research that stretches back 60 years. The university has an Engineering and Technology Department, which provides a vital and diverse technical base for promoting research, as well as the Emerging Media Initiative, which is responsible for ICT systems essential for research and education. Both have also built up collaborative relationships with the NII and other external research institutions. There is also the Frontier Science and Social Co-creation

Initiative, to which Prof. NAGAI is affiliated, which works to create research environments and support research activities.

"I believe that efforts to promote research and information systems being interlinked from the start was a strength. In that context, you could say that they developed naturally" (NAGAI).

There is also active dialogue among people in charge and decision-makers, and this "facilitates effective discussions, which is another of our strengths." However, the diverse faculties at the university each have their own intrinsic cultures and perspectives, and each has different attitudes and needs. Did these not become barriers?

"I think that is a challenge that is faced by all universities. I may not have experience across a large number of fields, but I have had experience working in fields that are extremely different from one another. I believe that has been beneficial for me" (NAGAI).

For example, organic chemistry, where there is a constant risk of accidents, and theoretical research, where ideas and wide-ranging discussions are crucial, differ in their attitudes concerning to what extent core lab hours should be observed. Additionally, daily routines at the School of Medicine and university hospital involve human lives.

"While I do not personally have experience with joint research with the School of Medicine, the current president of Kanazawa University is an alumnus of the school. The shared use of facilities at the School of Medicine is well-established with detailed rules in place. Meanwhile, the university's Faculty of Science does not impose such detailed rules, instead valuing free thinking" (NAGAI).

It is not the case that one approach is correct, while the other is wrong. People outside the School of Medicine who need to comply with the school's rules tend to feel that the rules are "too strict" or want "more freedom in using information." There is a reason for this high level of strictness, however.

"Expanding the School of Medicine's rules to other faculties would not be overly difficult. Doing the reverse, however, would be challenging. When establishing rules and policies, it may increase their likelihood of success if we keep in mind how they would work at the School of Medicine" (NAGAI).

The path toward a data management policy that clarifies the university's responsibilities

The Kanazawa University Academic Data Management Policy was established by the university in March 2022 and revised in July 2024 with a focus on effective data sharing. The focus is on clearly stipulating that "responsibility for data handling lies with the university," along with centralizing institutions for making decisions on sharing data other than academic papers and conference presentations, and understanding data sources in order to eliminate the acquisition of any illegal data. The details of the policy are quite in-depth, but as Prof. NAGAI puts it, "the policy is more like a constitution." Fleshing out the policy and going into specifics is the role of individual faculties and departments. But why is the policy even needed in the first place?

"The proper management of research data is crucial to promote open science, and a management policy is necessary for this management. The 6th Science, Technology, and Innovation Basic Plan has also set out the formulation of research data policies and data management as goals for each university" (NAGAI).

However, a policy "formulated merely because its formulation was mandated" is meaningless.

"Policies should be formulated as rules for promoting the management, utilization, and application of research data. This can also serve as a catalyst for realizing open science that really bears fruit."

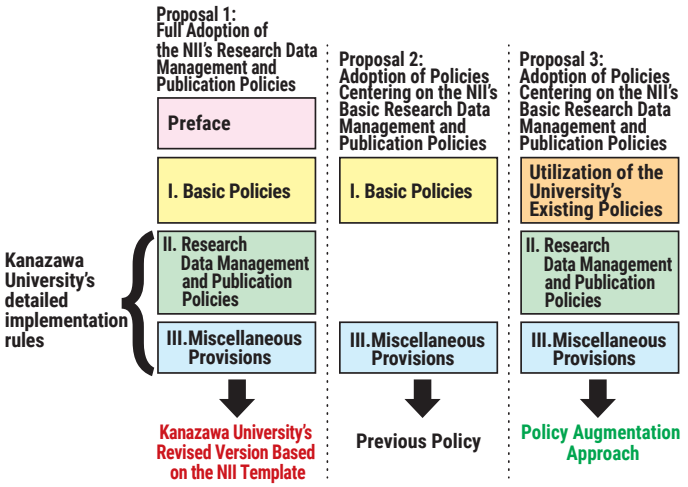
When Prof. NAGAI took up his post at Kanazawa University, a working group for policy formulation had already begun working on the policy. First, the policy was formulated as an "Academic Data Management Policy" in a form that included educational data in its content roughly similar to preceding policies at Kyoto University and Nagoya University.

"Using those precedents as a reference, we focused on making the policy as simple as possible. Trying to bite off more than you can chew with adding 'this and that' right from the start makes building a consensus difficult" (NAGAI).

When the university began implementing the policy, issues arose precisely because this educational data was included. Data can change even after it has been registered. And, the required response is not always the same. Thus, revisions to the policy began in fiscal 2023.

"The policy sets out important rules

Figure 1 The Hokuriku Research Data Infrastructure Consortium led by Kanazawa University is working on activities to establish rules for various institutions and develop and roll out systems to support these efforts, with the aim of supporting the management, publication, utilization, and application of academic data.
Created based on https://dri.w3.kanazawa-u.ac.jp/consortium/about.html and materials provided by Associate Professor NAGAI.



- The policies are like a constitution
- For revisions, NII templates are to be incorporated while ensuring legal compliance
  1. Ensure that the university takes responsibility (make the subject of every sentence "the university")
  2. Perspective on data publication
     ➡Everything other than academic papers and conference presentations to be referred to the data distribution committee
  3. Ascertaining data sources (eliminating illegal data acquisition)
     ➡Establishment of regulations for data exchanges in cases of personnel transfers or resignations, and joint research
- Future details will be stipulated in the detailed implementation rules. The "guidelines" serve as explanations for policies and detailed implementation rules

Figure 2 Formulation of Kanazawa University's Research Data Policy (Created based on materials provided by Associate Professor NAGAI)

*NII-RDP = research data management and publication policy https://rcos.nii.ac.jp/service/datapolicy/

for those within the organization; at the same time, it also serves as a public declaration by the university. The policy arrived at its current form after receiving guidance from attorneys to remove any ambiguities when read by people outside the university, such as amending all subjects to 'Kanazawa University' (Fig. 2). In terms of using NII's policy template as-is, while it was not quite that simple, the template proved extremely helpful as a sort of checklist" (NAGAI).

Of course, trialing the policy and examining the results within the university is no easy task.

"We are at the stage of having researchers with data to be published conduct case studies, as well as engaging in discussions with a diverse range of people, including administrative staff. Since data management itself is a new concept, it can be difficult to understand; we have therefore been taking pains in our efforts to work together to find terminology that everyone understands as meaning the same thing, as well as to ensure a shared vision regarding the outcomes of implementing the policy. We deliberately decided not to include in the policy any issues where consensus would be hard to reach. We should only include things in the policy that will definitely not change" (NAGAI).

Fine details and things that change depending on the situation can be specified in the implementation guidelines. Frequently needing to change the policy will also impact the university's credibility. The policy is just like a constitution.

Opportunities and challenges unique to Kanazawa and Hokuriku

In January 2024, an earthquake struck the Noto Peninsula in the Hokuriku region, and this was followed by heavy rains and flooding at the very northern tip of the peninsula in September. Alongside sympathizing with the suffering of those affected by these disasters, Kanazawa University also made strenuous efforts to digitize Kanazawa's ancient texts and cultural assets so that these can be passed on to future generations as data, too. During this process, "We were asked by the executive directors to save the data in GakuNin RDM before the information got scattered and lost. However, one challenge that arose was that using the system requires a certain degree of expertise. The second was that the number of participating institutions is limited, with very few private universities joining and no access for high school teachers. The third was that, due to the above, the system is difficult to use for urgent cases" (NAGAI).

Addressing these challenges is also homework for the NII.

"This is just the start for both the rules and systems. We are not yet at the stage of stable operation; rather, we are at the point where important case studies are emerging, and I hope to handle each one accurately, and if possible, swiftly" (NAGAI).

Advice for organizations working to develop a research data policy

"At the beginning, instead of trying to incorporate everything and anything, you should first aim to form institution-wide consensus. Even something small is fine, so creating a policy for the publication of data would be a good starting point. At the same time, you will need to gradually make progress on a variety of technical aspects, but starting up from the point of "What's GakuNin?" will require the support of the NII. That said, there are many things you won't be able to understand without first-hand experience, so a help list for those who know absolutely nothing may be necessary. The first steps are crucial.

When it comes time to revise the policy, I am sure that various universities will provide precedents from now on, so comparing each and copying the most easily applicable elements seems like a practical approach" (NAGAI).

Prof. NAGAI also has concerns regarding policy formulation.

"One worry I have is that it has become almost a default position for universities to delegate responsibility to faculties without taking any responsibility as an organization. Organizational restructuring and the like could make it unclear where responsibility lies. I believe that, to a certain extent, universities need to create policies that place responsibility on the university's shoulders" (NAGAI).

When universities consider formulating their own research data policies, they may gain additional insights by looking to regional consortia and other groups where pioneers like Prof. NAGAI are working.

Interview and writing: MIWA, Yoshiko/Photos: SUGISAKI, Kyoichi

NII NEWS TOPICS

Period: Aug. 1 (Thu.), 2024 to Dec. 31 (Wed.), 2024

More details about news items are available online.

www.nii.ac.jp/news/2024 (Japanese)

Facebook

www.facebook.com/jouhouken/

X (formerly, Twitter)

x.com/jouhouken

YouTube (audio will play)

www.youtube.com/user/jyouhougaku

Bit (NII Character)

X (formerly, Twitter)

x.com/NII_Bit

Send us your comments about NII Today.

www.nii.ac.jp/today/iken

Subscribe to our mailing list.

www.nii.ac.jp/mail/form

Read past issues of NII Today online (English).

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

The print version (Japanese version only) can be distributed to organizations such as school libraries and research institutions. Please email kouhou@nii.ac.jp to apply to receive or discontinue receiving the print version. It is not distributed to individuals, but the latest copies are always available to pick up from the 1st floor of the National Center of Sciences Building in Hitotsubashi, Chiyoda-ku.

News Releases

2024

Dec. 24 Fully open large language model with approximately 172 billion parameters (GPT-3 level) "llm-jp-3-172b-instruct3" released to public —Achieving performance beyond GPT-3.5

Nov. 28 CTC and NII promote joint research toward standardization of digital credentials at academic institutions
—Aiming to establish governance and technical standards for the digitization of degrees, academic histories, and student IDs—

Nov. 1 Sixth set of Line stickers featuring official NII mascot "Johoken Bit-kun" on sale

Oct. 21 Provision of trial version of CiNii Research's automatic translation function begins
—Enabling users who cannot speak Japanese to search Japanese academic papers—

Oct. 16 Fujitsu-led industry-academia consortium of nine organizations commences development of world's first disinformation countermeasure platform

Sep. 17 Full-scratch learning for large language model with approximately 172 billion parameters (GPT-3 level) underway and preview version of "LLM-jp-3 172B beta1" released
—The world's largest fully open model, including training data—

Sep. 2 Join a chemistry treasure hunt using informatics technology!
Competition involving using machine learning methods to discover catalysts for efficiently converting methane into ethane held

Awards

2024

Dec. 5 Paper by Professor ARAI, Noriko (Information and Society Research Division) and colleagues wins 30th Anniversary Paper Award awarded by the Association for Natural Language Processing

Dec. 4 Paper by Professor HASUO, Ichiro (Information Systems Architecture Science Research Division) and colleagues wins Best Paper Award at ICTAC 2024

Nov. 3 Professor Emeritus NEGISHI, Masamitsu conferred with "Order of the Sacred Treasure, Gold Rays with Neck Ribbon"

Oct. 2 Paper by former NII Project Researcher IINO, Nami (Principles of Informatics Research Division) and colleagues selected as "Specially Selected Paper" in Information Processing Society of Japan's Journal of Information Processing

Sep. 26 Joint research paper by Professor INOUE, Katsumi (Principles of Informatics Research Division) and colleagues receives Best Student Paper Award at LPNMR 2024

Sep. 11 Associate Professor BONO, Mayumi (Information and Society Research Division) receives Mitsubishi Foundation Humanities Research Grant

Sep. 5 Assistant Professor KOBAYASHI, Taisuke (Principles of Informatics Research Division) wins 38th Best Paper Award from Robotics Society of Japan

Aug. 15 Professor GOSHIMA, Masahiro (Information Systems Architecture Science Research Division) and colleagues receive Outstanding Research Award at xSIG 2024

Aug. 15 Mr. UCHIYAMA, Kazuhide (Goshima Lab, Graduate University for Advanced Studies, SOKENDAI) wins Poster Award at xSIG 2024

Aug. 6 Paper by former NII Project Researcher SAKAIDA, Rui (Future University Hakodate) and Associate Professor BONO, Mayumi (Information and Society Research Division) wins Excellent Field Research Report Award from Japanese Society for Qualitative Psychology

Aug. 2 Project Researcher OGAWA, Jun (Digital Content and Media Sciences Research Division/Kitamoto Lab) wins Yamashita SIG Research Award 2024 from Information Processing Society of Japan

← Opening Photo: Taken at the planned construction site of the Fukushima Institute for Research, Education and Innovation (F-REI), located right in front of JR Namie Station in Fukushima Prefecture. This place is approximately 5 km from the site of the Fukushima Daiichi Nuclear Power Plant. While evacuation orders remain in effect for part of the region, about 2,000 people are living in the town of Namie as of 2024. F-REI, where EMURA, Katsumi serves as Executive Director, was established in 2023 under the Act on Special Measures for the Reconstruction and Revitalization of Fukushima. The institute has begun operations at this site, aiming to become the "the Centre of excellence for creative restoration" at the forefront of innovation in science and technology, and industrial competitiveness.

Essay

Toward Creating Data Utilization Research Communities

TAURA, Kenjiro

Professor, Graduate School of Information Science and Technology, University of Tokyo
Executive Director and Vice President, University of Tokyo

As part of the Research Data Ecosystem Development Project, we are conducting a project titled "Use Case Creation." This involves soliciting and evaluating project proposals focused on the establishment, utilization, and application of research data, and providing support necessary for implementing the adopted projects. As of January 2025, 33 projects have been adopted. Details of the call for proposals can be found at https://www.nii.ac.jp/creded/use-cases.html, and a list of current and past projects is available at https://www.nii.ac.jp/creded/creded_result.html, so please take a look (Japanese only).

Those selected for adoption are asked to proactively make use of platforms for data aggregation, sharing, and utilization such as GakuNin RDM and mdx, to promote their projects. We also host symposiums featuring oral presentations, poster sessions, and panel discussions. These events provide opportunities for use case project participants to provide feedback on the overall Research Data Ecosystem Development Project and exchange ideas and information with one another. A Slack workspace has also been set up for use case projects, and is always available for questions regarding platforms like GakuNin RDM and mdx.

Supporting, promoting, and creating research with platforms for data-aggregation, -sharing, and -utilization like GakuNin RDM and mdx

Research utilizing information technology and information platforms has long been crucial in many fields. Aided by the rapid advancement of AI, however, the importance of aggregating and organizing data to make it searchable and utilizing that data effectively is growing for an even wider range of disciplines. In this project, by proactive using research platforms such as GakuNin RDM and mdx, which are offered to researchers nationwide, we are supporting the aggregation and organizing of research data, the sharing of data with joint researchers, and the promotion of large-volume data processing including projects that require significant computing resources, such as AI, in an integrated environment.

Permanent and upgraded national data platforms, and through this, enhanced research environments and research capacity

The approach of conducting research utilizing such platforms should be encouraged as standard practice for researchers, regardless of whether or not their research is conducted as part of the Use Case Creation Project. If it were a matter of course that these platforms are continually updated with the latest equipment and made accessible swiftly and cheaply nationwide (irrespective of affiliated institution), we could reduce the costs and time needed to develop separate research environments for each researcher, institution, or discipline, thereby improving research capacity. In addition, there are various requirements placed on information platforms, whether data platforms or utilization platforms, including linkages with experimental equipment, improving security, long-term storage of large volumes of data, and resistance to equipment breakdowns and natural disasters. Our goal is for the outcomes achieved through projects such as the Use Case Project will bring about a wider understanding of the need for such platforms, leading to a trend of creating next-generation platforms that respond to the various different requirements from different fields.

Interdisciplinary research connecting informatics with a variety of other fields

The informatics field is one that has traditionally provided tools (not only in the literal sense of a tool, but also common computational methods and algorithms used across disciplines) that can be used in many fields, and which has extensive ties to other disciplines. This project aims to further expand these connections to generate interdisciplinary research that links informatics with other fields, or links together multiple disciplines utilizing informatics.

Conclusion

The Use Case Creation Project accepts applications on an ongoing basis, with evaluations conducted every other month. In addition, we have devised measures for effectively supporting research, such as accepting applications spanning two fiscal years in a single submission. We encourage those researching areas that align with our objectives to consider submitting an application. Of course, you can utilize GakuNin RDM and mdx without needing to apply for the project, so please also consider taking advantage of these platforms.

Weaving Information into Knowledge
NII

National Institute of Informatics News: NII Today No. 104 Jan. 2025
Published by the Research Organization of Information and Systems, National Institute of Informatics
Address: National Center of Sciences, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430

For inquiries regarding this publication:
Publicity Team, Planning Division, General Affairs Department
Email: kouhou@nii.ac.jp

Publisher: KUROHASHI, Sadao
Editorial Committee Chair: ECHIZEN, Isao
Editorial Supervisor: YASUURA, Hiroto
Advisor: NAKANO, Keiichi
Editorial Committee Members: IKEHATA, Satoshi; KANEKO, Megumi; KOMIYAMA, Yusuke; MIZUNO, Takayuki; TAKEFUSA, Atsuko (alphabetical order)
Editor/Art Director: KISHIMOTO, Harue
Production/Printing: Nissei PR Co., Ltd.
Cover Illustration: ICHIMURA, Joe