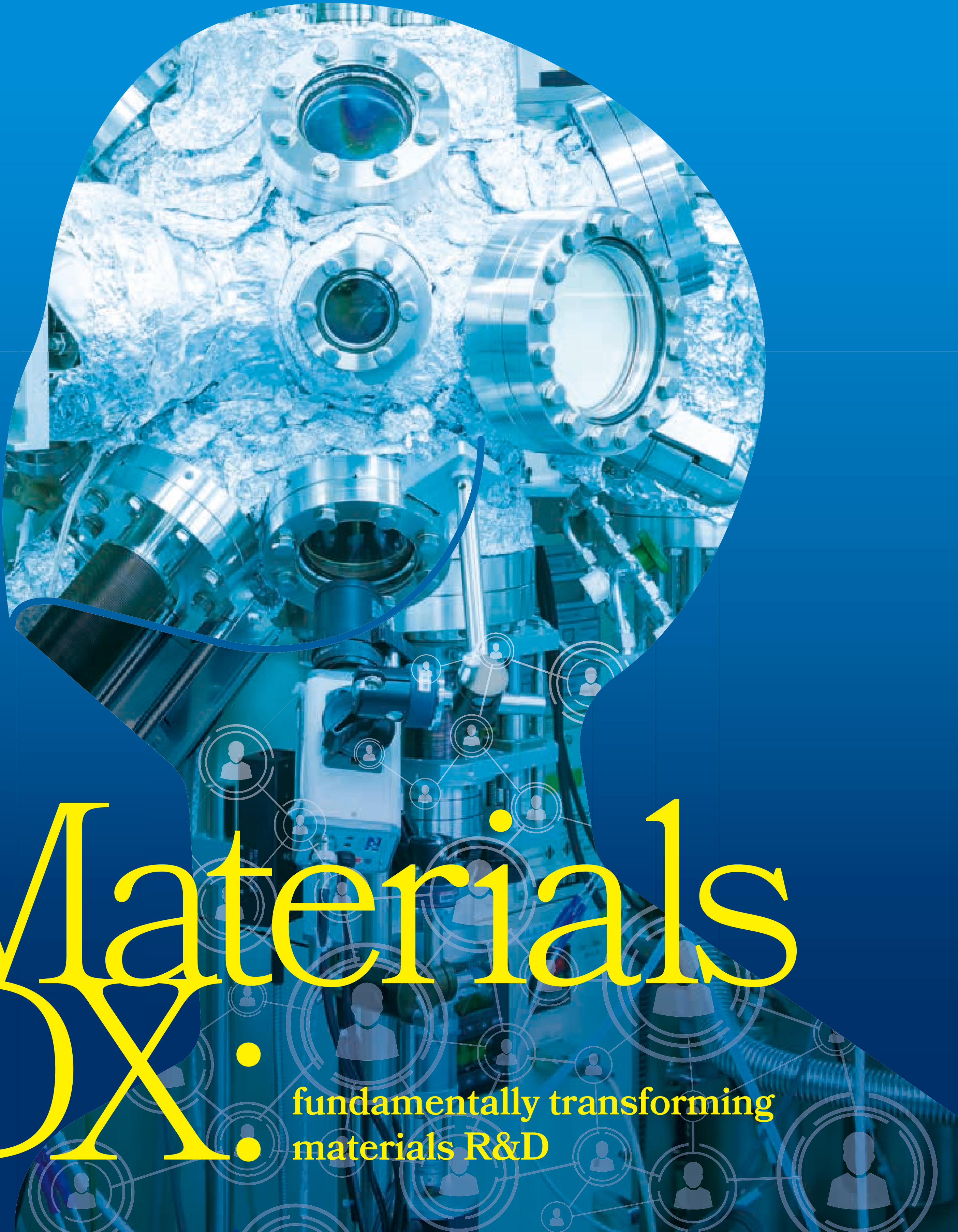
2021

NIMS NOW INTERNATIONAL

NATIONAL INSTITUTE FOR MATERIALS SCIENCE

No. 5



Materials DX:

fundamentally transforming materials R&D

Materials DX:

fundamentally transforming materials R&D

As part of its materials informatics initiative, NIMS has incorporated the use of AI and data science into its efforts to discover new materials with desirable functions and characteristics. In addition, this initiative has been promoting the introduction of digital transformation (DX) into materials development.

The availability of systems enabling efficient data sharing among materials researchers from different sectors and organizations was found to be a vital aspect of implementing DX.

A vast amount of diverse materials data is produced across Japan on a daily basis. Making this data efficiently and adequately sharable among all of the members of the materials R&D community will require the development of open platforms accessible to all material-related sectors and organizations. NIMS has developed efficient data collection and utilization techniques, and operated an open facility to make its array of state-of-the-art equipment accessible to external researchers. Using these experiences and resources, NIMS has just begun leading nationwide efforts to build so-called materials DX platforms.

Adopting DX may transform not only the ways in which materials are developed but also the fundamental approach to materials research. This NIMS NOW issue spotlights DX in materials R&D.

Masahiko Demura

Director, Research and Services Division of
Materials Data and Integrated System (MaDIS)
National Institute for Materials Science (NIMS)

×

Motoi Eto

Director, Materials Science and Nanotechnology Division
Research Promotion Bureau
Ministry of Education, Culture, Sports, Science and Technology (MEXT)

Special talk

Benefits of implementing digital transformation (DX) into materials R&D

Since FY2020, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) has been promoting coordinated efforts to develop "materials DX platforms"—nationwide materials research platforms designed to facilitate strategic materials data collection, storage and distribution by the industrial, academic and public sectors. This framework is intended to encourage efficient and continuous data production and sharing by these sectors. NIMS is building a core database system for these platforms by upgrading conventional data collection and utilization methods.

What are the ideal forms of materials DX platforms for Japan's materials R&D? What is the role of NIMS in the development of these platforms? Masahiko Demura (MaDIS Director, NIMS) and Motoi Eto (Director of the MEXT Research Promotion Bureau) discussed the significance of developing materials DX platforms.

NIMS' effort to develop materials DX platforms

Eto: The approach of analyzing big data using AI is becoming increasingly popular in materials development around the world. The world-class materials databases NIMS operates have been serving as a vital resource for the materials industry.

You've been leading a research project entitled "Establishment of the inverse design MI basis for advanced structural materials and processing" under the "Materials integration for revolutionary design systems for structural materials" category of the SIP program* led by the Cabinet Office. I understand that the aim of this project is not only to further enhance NIMS databases and increase data production efficiency but also to implement DX into materials R&D to enable the analysis of vast amounts of complex materials data.

* SIP: Cross-ministerial Strategic Innovation Promotion Program

Demura: NIMS has been operating its materials databases for many years. We began constructing a next-generation materials data platform in 2017 with three focuses: development of adequate data storage facilities, efficient data collection, and utilization methods. We then started building a core database system in FY 2020 as part of the MEXT-led effort to develop materials DX platforms. Moreover, NIMS was selected as the central hub of the Advanced Research Infrastructure for Materials and Nanotechnology in Japan —"creation of structures for the accumulation and use of data in conjunction with the use of advanced shared facilities"—in April 2021. Analyzing big data using AI is becoming a common approach in materials development as a way of expediting development processes. The implementation of DX is vital in making research data collected by the industrial, academic, and public sectors more widely and easily accessible.

Eto: Some of the SIP projects I'm directing also aim to increase data collection and utilization efficiency. In one such project, "Smart Logistics Service," we're aiming to maximize the delivery efficiency of logistic companies by collecting commodity distribution data from them using IoT technologies and processing this data using cloud servers.

Different logistics companies transport goods to the same areas, and their transportation efficiencies would significantly increase if they simply shared the same freight trucks instead of using their own trucks. Field experiments conducted in an SIP project demonstrated that this practice could greatly increase transportation efficiency in terms of various parameters, including the total number of working hours, gas expenses, and CO2 emissions. Although logistics companies generally do not share commodity distribution information, they are beginning to consider organizing and sharing their data to overcome a common challenge they are all facing: a serious labor shortage.

Demura: I believe that the very essence of DX is to open up various barriers existing in a society. An increasing number of companies in the materials industry have been showing interest in sharing data in recent years. The implementation of DX may help open up the barriers between companies.

A large number of materials researchers are working at NIMS. They all have different areas of expertise (e.g., physics, chemistry, metallurgy and mechanical engineering) so that they carry out research on a wide range of materials used for different purposes, including functional and structural materials. Their areas of expertise have recently been further specialized, making it more difficult for them to follow the research of others. This trend may deepen divisions between different research groups, making collaborative interactions more challenging. We believe that implementing DX at NIMS would be very beneficial in helping remove barriers between research groups with different specialties.

Eto: Japan's materials industry remains strong in the global market. However, Japanese materials manufacturers have a strong sense of crisis in maintaining their competitiveness over their competitors overseas. They are probably aware of the urgent need to share the R&D data they have accumulated and conduct collaborative research to accelerate R&D in order to maintain their global competitiveness.

I really hope that implementing DX will help break down barriers between the industrial, academic and public sectors, and between researchers, and help strengthen the international competitiveness of Japan's materials industry. The national government intends to fully support this initiative.

"I want to build materials DX platforms capable of facilitating active communication between researchers with different expertise to generate new ideas." — Masahiko Demura

Data structure is the researcher's view of the research itself

Demura: You mentioned earlier that logistics companies are reluctant to share commodity distribution data. I suspect that this reluctance is not related to company policies but can rather be attributed to a lack of a common data format they can use to exchange data. In other words, if a common data format were available to them, they would be willing to share data to maximize profits by optimizing their logistics operations. Therefore, establishing common data formats would be an initial breakthrough in DX implementation.

Eto: I completely agree. During the smart logistics project, I described above, we created a rule that all delivery vehicles owned by different logistics companies would be identified using vehicle identification numbers assigned to them by the Transportation Bureau. We also converted different commodity names used by different logistics companies into consistent names during the process of constructing DX platforms for these companies. As a result, we were able to experimentally demonstrate improved transportation efficiency and the advantages of data sharing to them in a tangible manner.

Demura: We are working on to achieve similar goals, including the establishment of common data formats and the development of data conversion methods.

Eto: Establishing common data formats is an important process in the development of materials DX platforms. It is also important for background information associated with research data (i.e., data structure) to be accessible to platform users. Otherwise, the platforms would not be very useful to many researchers. What are your thoughts on this?

Demura: I have the same opinion. I recognize three layers of information in every research output. First is the data itself. Research results are based on data, so data is fundamental. Secondly, there is the layer of data structure, which includes information on the methods used to collect data, the authors' reasons for making specific measurements and the ways in which raw data was processed into the data presented in the publication. In the smart logistics project discussed above, data formats fall into this category. Other types of data structure may include date/time of data collection, names of researchers involved, research objectives, names of materials researched, material compositions and methods used. They should be listed in an orderly and consistent manner in tables. Data can only be measured if there is a data structure. In a sense, the data structure is the researcher's view of the research itself. The third layer is called models, contains mathematical models used to describe physical relationships between different parameters measured.

Eto: I see. The three layers of information consist of data itself, data structure, and models, correct?

Demura: That's right. These three types of information are provided in every research paper. However, these papers—focused on clearly conveying the authors' points to readers—do not clearly provide these three types of information in an organized and consistent manner. It is therefore desirable for data providers to organize their data into the machine-readable, three-layer format before submitting papers. This practice would further improve the operational efficiency of materials DX platforms. I envision the use of these-layer platforms not only making materials research more efficient but also serving as a new communication framework for researchers.

Designing ideal materials DX platforms

Demura: Data curation is the organization and integration of data collected from various sources. One important part of the curation process is constructing appropriate data tables to organize research data and data structure collected from various research outputs so that they can be accurately interpreted. This process is challenging because the types of data structure generated vary from one research output to another. This table structure is similar to the structure of Excel spreadsheets. Many insights can be gained from data structure generated by other researchers. As new scientific insights replace old ones, the types of data structure generated are likely to change. We need to find ways to properly digitize dynamically changing data structure. This is an important aspect of adopting DX in basic research.

Eto: We must create materials DX platforms capable of facilitating—not hindering—researchers' creative activities and mutual communications. We should expect that the types of data structure provided by researchers will vary widely and change over time due to differences in researchers' working styles and cultural backgrounds and due to changing social needs.

Demura: Studying other researchers' data structure gives platform users a chance to gain the original researchers' insights on their research subjects and to learn their manners of expression, principles and philosophies. In addition, examining data structure derived from multiple research outputs allows users to compare and contrast it and recognize conceptual hierarchies (e.g., concept A is superordinate to concept B). These activities, including the verification of new concepts, are purely scientific in nature.

Traditionally, communication among researchers has been via scientific publications and research presentations at scientific meetings. When I was young, we were subscribing to monthly scientific journals. While reading them, I often skimmed through research articles that were not directly related to my area of research, allowing me to familiarize myself with a broader range of research topics. In addition, while there were fewer scientific meetings held at the time than today, they generally covered a broader range of disciplines. Attending only one annual meeting allowed me to widen my perspective within materials science. The number of research publications has greatly increased today and most researchers nowadays regularly search for research articles of interest through the internet. Scientific meetings have also become highly specialized, greatly limiting researchers' opportunities to learn about different areas of research and reinforcing their tendency to devote themselves to their highly specialized areas of research. This trend may throw them out of balance in terms of learning experience. I hope that implementing DX will improve these circumstances for researchers.

Eto: What you just described is an example of how DX may break down barriers between researchers and between research groups with different expertise.

Demura: Exactly. I think the inclusion of data structure will add value to materials DX platforms. The collection and organization of data structure (i.e., curation of data structure) is expected to enable researchers to utilize data from wider areas of research, potentially increasing their research efficiency, giving them new insights and promoting the generation of new scientific concepts. This is why I've been advocating for the importance of data structure curation even more than the curation of research data itself. I plan to spend the next 10 years developing materials DX platforms capable of facilitating active communications between researchers with different expertise and the generation of new concepts.

"I hope implementing DX will help break down barriers between the industrial, academic and public sectors and between researchers and help strengthen the international competitiveness of Japan's materials industry." — Motoi Eto

Eto: Roughly speaking, I previously understood materials DX platforms to be a mere framework for increasing the efficiency of data collection and developing new databases, thereby facilitating data sharing. I now realized that these platforms are expected to play a much greater role: their use may potentially promote communication between researchers, broaden researchers' perspectives, stimulate their creativity and increase their chances of making scientific discoveries.

Demura: That's right. Concurrent with these efforts, we're working with our collaborators to develop an Advanced Research Infrastructure for Materials and Nanotechnology—a framework to collect research data from shared facilities at universities and other organizations across Japan. This infrastructure would play a vital role in materials DX platforms.

Eto: After talking with you, I feel that implementing DX may even fundamentally change the ways research institutions operate. This DX initiative may potentially remove not only barriers between researchers but also temporal and spatial constraints. As a result, researchers may gain greater access to external resources in addition to the resources available within the specific research organizations which they belong to. I enjoyed this in-depth discussion on DX with you. Thank you very much.

Demura: Thank you very much.

(By Kumi Yamada)

Launch of DICE, materials data platform enabling efficient data collection and sharing

Mikiko Tanifuji
Managing Director,
Materials Data Platform Center (DPFC)
Research and Services Division
of the Materials Data and Integrated System

NIMS released DICE, its materials data platform, in June 2020 with the aim of aggregating high-quality data and creating user-friendly databases. These efforts are expected to generate a virtuous cycle of data collection and sharing. We asked DPFC Managing Director, Mikiko Tanifuji about DICE, a core component of the materials DX platforms under development.

DICE provides resources needed to initiate and complete data-driven research

The innovative DICE system has two key functions: it supports the initial stages of data-driven research by efficiently collecting and aggregating data generated by researchers and then facilitates the completion of the research by processing collected data into AI analysis-compatible formats and offering analytical tools to researchers. DICE collects a wide range of data, including researcher-generated data (e.g., experimental/measurement data and computational data), data from publications (e.g., research articles and patent documents) and external databases resulting from collaborative research between NIMS and other organizations. Research data is generated in many different formats and some can only be read by proprietary software, making it incompatible with machine learning processes. DICE has tools able to convert various data formats into machine learning-compatible formats. Moreover, DICE offers a number of convenient applications, including simple analytical software, imaging systems and advanced systems capable of predicting materials performance using big data. DICE is therefore truly a comprehensive materials data platform capable of facilitating data creation, utilization and publication.

Tanifuji plans to make DICE more multifunctional and user-friendly in accordance with the FAIR principle. She says, "In the context of data services, FAIR stands for findable, accessible, interoperable and reusable. To promote the public use of our databases, we need to develop a system that allows users to easily find and access the types of data and databases they're looking for. In addition, when private companies want to link their databases for their analyses, we'll provide them with a linkage solution and deal with data format compatibility issues. We should therefore store our data in forms conducive to meeting researchers' diverse needs. We've been standardizing our databases in line with the FAIR principle and the global trend to treat data as a common asset for data-driven sciences."

DPFC is working to make the DICE platform user-friendly in various ways. These include: assigning permanent identifiers to individual datasets to make them easier for users to find, improving user accessibility to databases by adopting open communication protocols, using a common description language to facilitate interoperation, and issuing formal data reuse licenses. While user-friendly services are important, security issues also need to be addressed as these services directly impact Japan's international competitiveness in materials research. Tanifuji says, "Data services for the academic community should be user-friendly and open in accordance with the FAIR principle. However, they also need to be secure for use by private companies to contribute to Japan's international industrial competitiveness. We'll maintain a careful balance between these two imperatives while operating DICE."

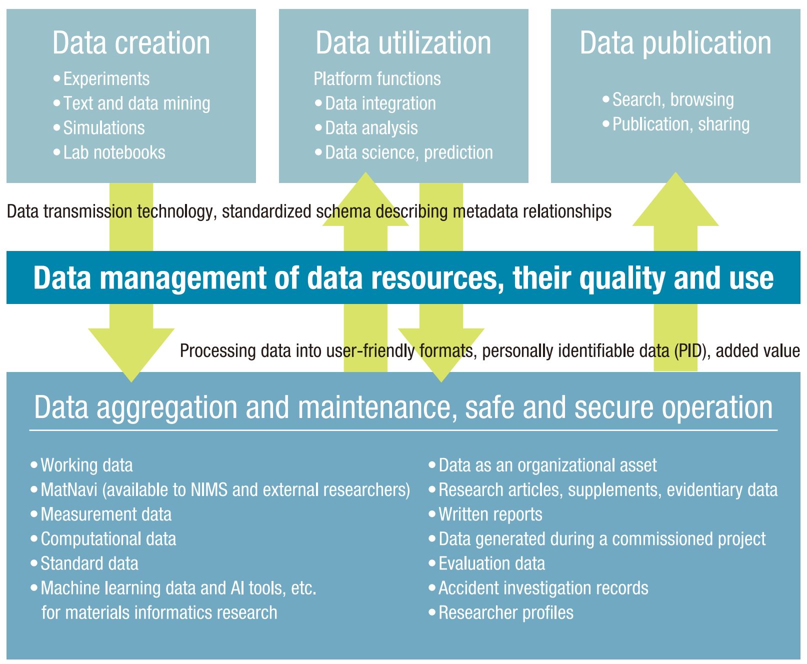


Figure. Key functions of the DICE materials data platform

Four key DICE resources

1 MatNavi: world-class materials database

MatNavi, released to the public in 2003, is one of the largest materials databases in the world. It consists of more than a dozen sub-databases, including the PoLyInfo polymer database, the AtomWork and AtomWork-Adv inorganic materials databases, the Kinzoku metallic materials database and NIMS' structural materials databases (e.g., creep and fatigue data sheets). In addition, MatNavi offers application systems, including the SurfSeg metal segregation prediction system and the InterChemBond interfacial chemical bonding prediction system. Tanifuji states that, "To promote the use of the DICE materials data platform, we must commit to offering high-quality databases. MatNavi contains expert-reviewed, superb-quality data which has been collected and aggregated for nearly 20 years, making it one of the word's best materials databases, in terms of both quality and quantity."

There is one major issue with MatNavi, however: the individual MatNavi databases are operated independently and are not cross-searchable.

In relation to this issue, Tanifuji says, "Back in 2003 or so, digitized data available for use by the public free of charge were rare. I recall that data and description formats and data structures at the time were largely inconsistent between different types of materials data (e.g., electronic and catalytic materials) and different areas of research, even within a single inorganic materials database. Active discussions on data compatibility among researchers, between projects, and on the assignment of permanent identifiers for individual datasets to ease data searches, started only within the last decade or so. Today, the FAIR concept is widely accepted by the global researcher community, and data is now generally considered to be computer-processable digital information."

DPFC has been assigning permanent identifiers to individual datasets based on the FAIR principle. In addition, an increasing number of private companies have been requesting that DPFC link MatNavi to their confidential databases. To satisfy this demand, DPFC is developing application programming interfaces (APIs) that can be used to connect different databases. Tanifuji comments on this development as follows, "We've already developed an API compatible with the PoLyInfo polymer database and machine learning applications. We're currently conducting technical studies to verify that the method of authorizing the use of the API works properly in conjunction with various joint research schemes. We plan to release the API for public use after verification is completed."

2 RDE: system to automate research routines

Handwritten lab notebooks were once the only way to record research data. Today, significant progress has been made in the development of technologies capable of real-time aggregation of measurement data onto data servers and automatic analysis and prediction. The Research Data Express (RDE) system is the first such technology close to practical use (see page 11).

The RDE system is capable of automatically and securely transmitting any type of data generated at research labs to a data server, including data generated by lab equipment and information recorded on electronic lab notebooks, such as materials used and experimental conditions. However, most binary raw data generated by lab equipment can only be read by proprietary software. To convert equipment-specific terminology and data formats into common formats, DPFC has developed a translation tool called M-DaC. In addition to this, it is developing other tools capable of automatically performing simple data analysis immediately after measurements are taken (see pp. 11–13 for details on these systems and tools). The MDR repository (described below) will allow its users to publish data stored in it and cite it as a reference in their research publications.

According to Tanifuji, "Being a materials development organization itself, NIMS knows the needs of other materials development laboratories: efficient and continuous collection of high-quality data. RDE is a comprehensive data system capable of executing a series of data-related tasks, from routine data collection at research labs to transforming data into graphs and displaying them in a simple web view for easy comprehension and analysis. We plan to use the RDE system to enhance the efficiency and productivity of research labs connected to it. We present these labs as 'smart laboratories' at NIMS."

3 MDR: FAIR principle-based data repository

A data repository offers online data storage, sharing and publication services. NIMS has been operating its materials data repository (MDR) which allows its users to share and publish various types of data—including experimental, measurement and analytical data—and makes analytical tools and programs available to them. The MDR offers two types of data sharing: limited data sharing by researchers participating in the same joint research (MDR Closed) and public, open data sharing in which DOIs are assigned to datasets (MDR Open). In addition, data collection methods are designed to ensure that data stored in the MDR is properly structured for AI analysis.

As Tanifuji states, "Published data needs to be structured differently for different purposes, such as for machine learning analysis and for material sample synthesis. For this reason, we have data specialists, including data curators and data architects, who are skilled in converting data structures to meet various needs before registering them. With their support, we have developed a user-friendly MDR repository based on FAIR principles by adopting globally standardized data communication protocols and metadata description formats."

4 MInt: system to accelerate collaboration between industry, academia and the public sector

MInt (Materials Integration by Network Technology) is a system capable of making comprehensive predictions about various aspects of metal-based structural materials, including fabrication processes, physical properties and performance. This system is equipped with various types of modules—AI-based computational tools that exploit material engineering theories and empirical laws. These modules can be combined in any way to create a workflow for a specific objective. For example, suppose that you want to accurately estimate the creep rupture lives of welded joints made of heat-resistant steel used at high temperatures under high pressure in thermal power plants, and thereby determine optimum welding conditions. While it would take thousands of hours to solve such problems experimentally, the MInt system can computationally solve them in only a few hours.

The MInt concept was proposed and verified during the innovative structural materials development project (2014–2018) supported by the Cabinet Office-led, Cross-ministerial, Strategic Innovation Promotion Program (SIP). Subsequently, the materials integration consortium was established in December 2020 to promote the use of the MInt system by the industrial, academic and public sectors. This was one of the achievements of the ongoing SIP project started in 2018 entitled, "Materials integration for revolutionary design system of structural materials." As of October 2021, seven private companies and 16 academic labs and centers have joined the consortium and are working at their respective research bases to expedite their R&D using the MInt system. In addition, the industrial and academic sectors are jointly carrying out new problem-solving projects to add new modules and workflows to the MInt system. These efforts are expected to enhance the system's applicability. "I hope that these projects will promote cross-sectoral materials data utilization, which will then further improve the DICE platform, creating a virtuous cycle of data utilization and platform enhancement," Tanifuji says.

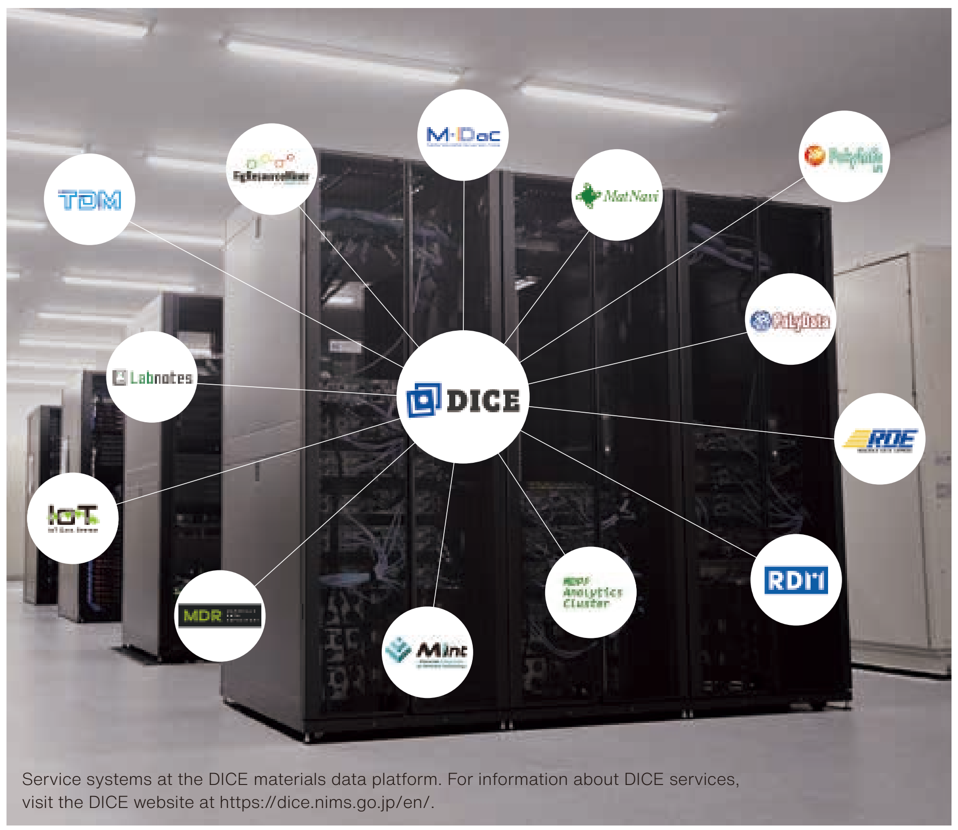
Creating a culture that encourages data sharing

The DICE materials data platform will ideally generate a virtuous cycle; that is, researchers will be able to securely submit their data, and user-friendly databases will be available for use by researchers. These attractive features will encourage DICE users to continuously submit high-quality data to the platform. According to Tanifuji, this cycle can be accelerated by developing systems that reward researchers who actively share their data, thus creating a scientific culture favorable to data sharing.

Tanifuji states that, "Most Japanese researchers seem to have a strong sense of attachment to their own data. To change this 'own' concept, I would like to deliver data scientists and data architects who advocate data-sharing to be actively involved in materials research. Their stimulating initiatives may increase materials researchers' interest in data sharing, and promote their more active use of data platforms. DICE also needs to evolve to become more attractive to the materials science community. As the DICE platform grows, I expect younger researchers will be more willing than veteran researchers to accept and incorporate new approaches to materials research—including data science and materials informatics. I believe this shift will lead to the development of a scientific culture supportive of data sharing. We therefore hope to develop DICE into a high-quality data platform attractive to next-generation researchers."

The DICE materials data platform based on FAIR principles is expected to accelerate the adoption of machine-learning techniques and data-driven research in materials science, by encouraging researchers to share their experimental and computational data with others, and to process raw data into AI-friendly formats.

(By Takeshi Komori)



Service systems at the DICE materials data platform. For information about DICE services, visit the DICE website at https://dice.nims.go.jp/en/.

Revolutionizing materials research labs

RDE system enables real-time, automatic data collection and analysis

Hideki Yoshikawa
Deputy Managing Director,
Materials Data Platform Center,
Research and Services Division of
the Materials Data and Integrated System

The amount and quality of available data influences the accuracy and effectiveness of data-driven research. NIMS has developed an RDE (Research Data Express) system capable of collecting the huge amounts of data generated daily at research laboratories and aggregating it into databases. The introduction of the RDE system will not only promote adoption of DX in materials research but may also transform the way in which research is carried out. Hideki Yoshikawa, Deputy Director of the Materials Data and Integrated System, led the development of the RDE system. We asked him about it.

Safely transmitting data from older equipment

In principle, researchers collect their own material-related data using the lab equipment they have access to. An automated, real-time data collection system would greatly facilitate the efficient gathering of data generated by individual researchers and its aggregation into central databases. Introducing such a system would also accelerate DX at research labs. Data generated by lab equipment is usually stored in the personal computers (PCs) directly connected to the equipment. Sending stored data from these PCs to a data server would seem easy to achieve. However, research labs have a unique issue that makes this process challenging: never updated operating systems (OSs).

Because many lab instruments are very expensive, they are commonly used for many years. They are typically controlled by PCs dedicated specifically to this purpose. The OSs installed on these PCs tend never to be updated because of the fear that updating them could cause a control function failure. In addition, PCs with dated OSs cannot be connected to a network due to a lack of protection from security threats. Users of these lab instruments therefore transfer measurement data from equipment-controlling PCs to general-purpose PCs via data storage devices, such as USB flash drives.

Yoshikawa and his colleagues explored ways of safely connecting equipment-controlling PCs with dated OSs to a network. As a result, they developed a data transfer system using an IoT security device that enables one-way network communication from an equipment-controlling PC to a data server. Use of this device (see the photo on p. 12) protects the PC from cyberattacks while preventing the OS from being automatically updated. "Simply connecting this device to the LAN port of an equipment-controlling PC enables one-way communication," Yoshikawa said. "Users need only save measurement data in a designated folder on their equipment-controlling PCs, and the data will automatically be transmitted to the data server. Use of this device has enabled us to automatically collect measurement data from various lab instruments without security issues."

Translating equipment-specific languages

In addition to automating data collection, other issues need to be resolved before practical databases can be created. The languages and formats used to generate output data vary between different types of lab instruments and different

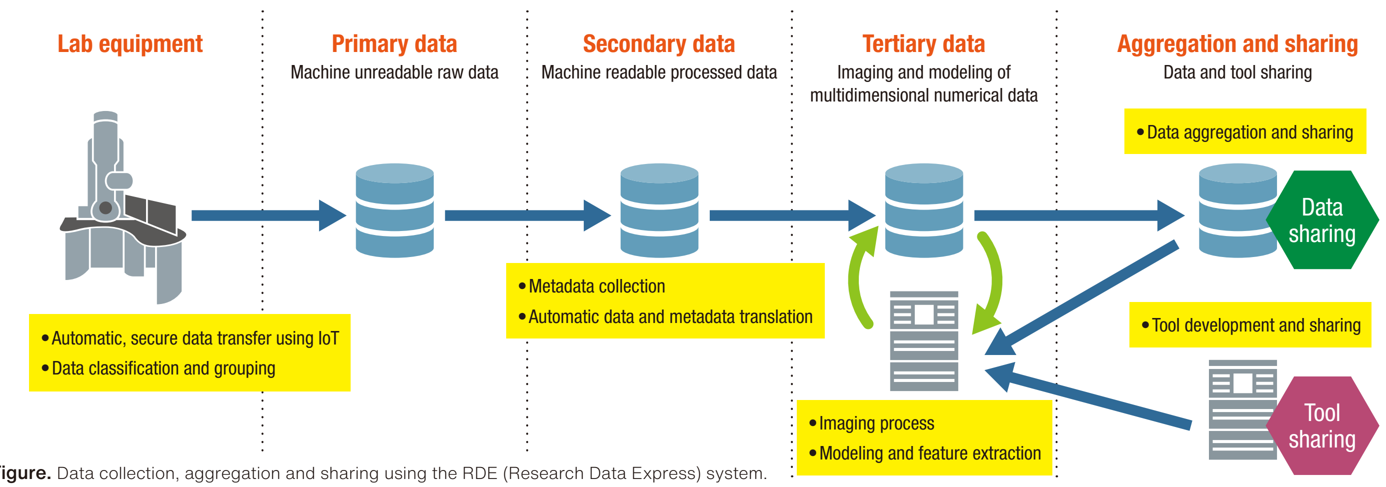


Figure. Data collection, aggregation and sharing using the RDE (Research Data Express) system.

IoT security device enabling one-way communication

manufacturers. Moreover, some measurement data can be read only by the lab equipment's proprietary software. To address this issue, Yoshikawa's group has developed a translation tool called M-DaC with the cooperation of equipment manufacturers. M-DaC can convert data formats readable only by proprietary software into commonly used formats and translate various languages and formats used by different instruments into highly readable, consistent forms, allowing aggregation of AI-compatible data into data platforms.

Output data files generated by lab equipment also include metadata (e.g., information on experimental conditions and the type of equipment used) in addition to numerical measurement data. Metadata is needed to enable an AI to search, sort and analyze measurement data. "Metadata that cannot be automatically extracted from data files needs to be manually entered, which may cause entry errors and missing metadata," Yoshikawa said. "To minimize these human errors, we developed a user interface HTML template which allows users to select numerical values or string of characters from a pulldown menu and warns them of entry errors and abnormal value entries."

In addition to collecting data, the RDE system can also process and analyze data efficiently. For example, it is difficult for users to evaluate the quality of numerical data merely by inspecting it visually. The system can transform it into graphs and images in real time, helping researchers determine its quality. Moreover, researchers may have large amounts of complex data (e.g., spectral data generated by a spectrometer and microscope image data) whose quality is difficult to evaluate quickly. Yoshikawa's group is developing a data compression technique by which large amounts of complex data can be reduced to mathematical model parameters, allowing speedy automatic determination of its quality. "I also plan to develop machine learning and AI-based data analysis tools capable of facilitating the development of higher performance materials," Yoshikawa said. "Advice from materials researchers is vital for this endeavor."

Promoting data-driven research at private companies

The MEXT-funded Advanced Research Infrastructure for Materials and Nanotechnology project (see p. 14) was launched in FY2021. This project aims to aggregate materials data generated by national research organizations, universities, private companies and other research organizations into materials data platforms. Although private companies normally keep their research data confidential, participants in this project intend to encourage companies planning to use advanced materials research facilities to actively share their data. Another objective of this project is to expand the applicability of the M-DaC translation software to a wider range of lab instruments with the support of MEXT. Negotiations with lab equipment manufacturers are underway to obtain their permission to translate their unique terminology and formats used in output data files into sharable forms.

"The number of private companies that have recognized the benefits of sharing basic research data is increasing amid intensifying international competition in materials development," Yoshikawa said. "I want more companies to realize the benefits of data sharing by enhancing our databases and making them more user-friendly. To achieve this, seamlessly linking a larger number of lab instruments to the RDE system is vital."

Many researchers are expected to benefit from the RDE system, which is capable of automatically collecting measurement data and collecting and translating metadata and making it sharable. According to Yoshikawa, another benefit of this system is that collected data will be shared by researchers of diverse backgrounds. This may enable them to study data unrelated to their areas of expertise from unconventional angles, potentially leading to new insights and discoveries.

"For example, superconducting material researchers may not be interested in the discovery of a novel material if it doesn't have superconductive properties," Yoshikawa said. "On the other hand, dielectric material researchers may find it extremely intriguing—an example of serendipity. Because data sharing will give researchers opportunities to study other researchers' experimental data in addition to their own, this initiative may fundamentally change the way in which research is conducted."

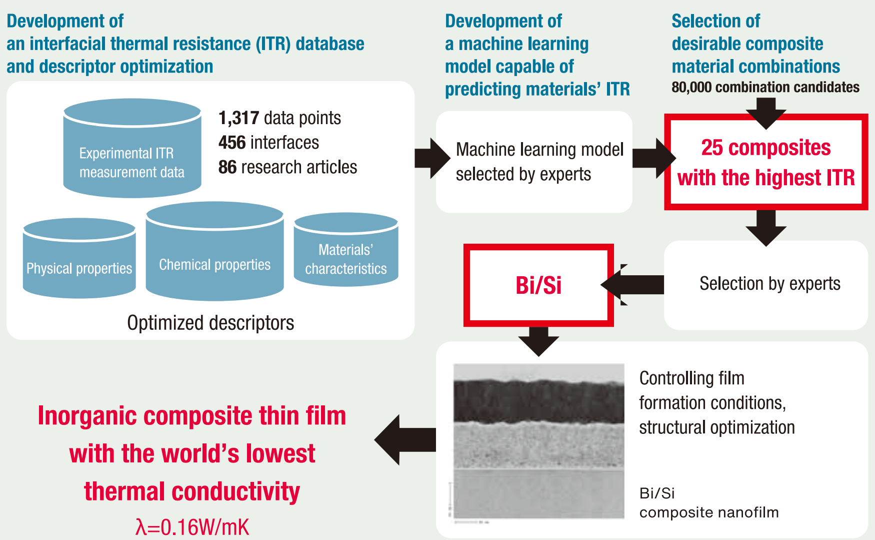
(by Kumi Yamada)

Exploration of new materials using abundant, high-quality data ①

Design and development of a nanocomposite material with the world's lowest thermal conductivity

Yibin Xu
Deputy Director,
Research and Services Division of the Materials Data and Integrated System

High-performance thermal insulation materials are vital to increasing the energy efficiency of various systems. Xu used AI techniques to analyze existing data of interfacial thermal resistance and a number of parameters extracted from more than 80 research articles and materials databases. Machine learning models were used to first predict interfacial thermal resistance with the parameters and then select 25 materials with the highest thermal resistance out of approximately 80,000 candidate materials. Xu's group ultimately chose Bi/Si after considering various factors, including ease of synthesis. The thermal insulation properties of Bi/Si thin films were optimized by adjusting their film synthesis conditions. A thin film in which Bi nanograins dispersed in amorphous Si matrix was the final material product which achieved the world's lowest thermal conductivity for inorganic composites.

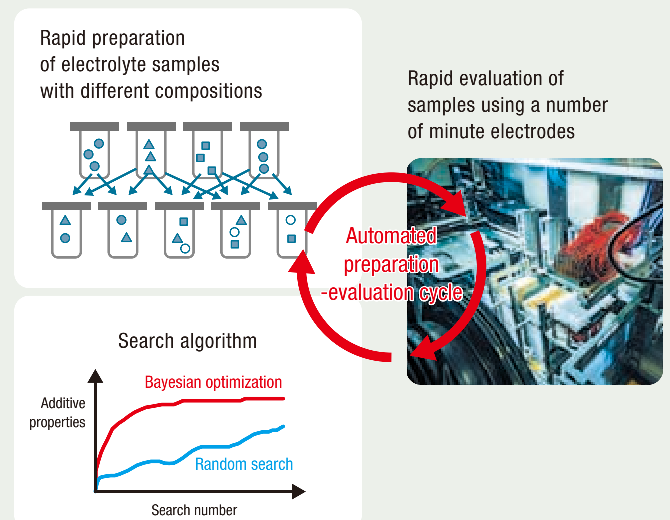


Exploration of new materials using abundant, high-quality data ②

High throughput electrolyte search system

Shoichi Matsuda
Senior Researcher, Rechargeable Battery Materials Group
Center for Green Research on Energy and Environmental Materials

International efforts to reduce CO_2 emissions are underway through the development and popularization of green technologies, such as electric vehicles (EVs). High-performance rechargeable batteries are indispensable EV components. In addition to the composition of their positive and negative electrodes, electrolytes influence the performance of rechargeable batteries. The number of potentially effective electrolytes is enormous, and the conventional, manual, repeated trial-and-error approach to identifying high-performance electrolytes has become virtually impractical. To address this issue, Matsuda developed an automated electrolyte search system in which electrolyte samples are prepared by robotic arms and their performance is evaluated automatically. Evaluation data is reviewed to adjust electrolyte search algorithms which are then used to prepare new samples composed of selected candidate electrolytes. This preparation-evaluation cycle is repeated automatically. This system is capable of evaluating 400 samples a day. In the project to identify highly functional electrolytes for lithium electrode, this system was able to select the five most promising electrolyte components out of 16 compounds and determine the composition of a new, effective electrolyte (see also the Vol. 20, No. 2 issue of NIMS NOW).



Development of Advanced Research Infrastructure for Materials and Nanotechnology to collect data nationwide

Yasuo Koide
Senior NIMS Scientist with Special Missions
Organizing Director,
Advanced Research Infrastructure
for Materials and Nanotechnology Project

The goal of the advanced materials research infrastructure project is to collect, aggregate and share high-quality data generated across Japan. NIMS will lead this project by leveraging its 20 years of equipment sharing know-how.

The Nanotechnology Platform Japan (NPJ) program which launched in 2012 is a network of shared advanced facilities and equipment owned by universities and national research organizations. These facilities and equipment have been used more than 3,000 times annually by a total of more than 20,000 researchers. The Advanced Research Infrastructure for Materials and Nanotechnology project was newly launched to collect, format and aggregate data generated nationwide, upgrade and expand existing databases by taking full advantage of existing shared facilities, human resources, know-how and real-time measurement data collection techniques (see p. 11–13).

"This is the first ever project of this kind in the world—including the United States, despite its highly developed equipment sharing networks," said Yasuo Koide, the organizing director of this project.

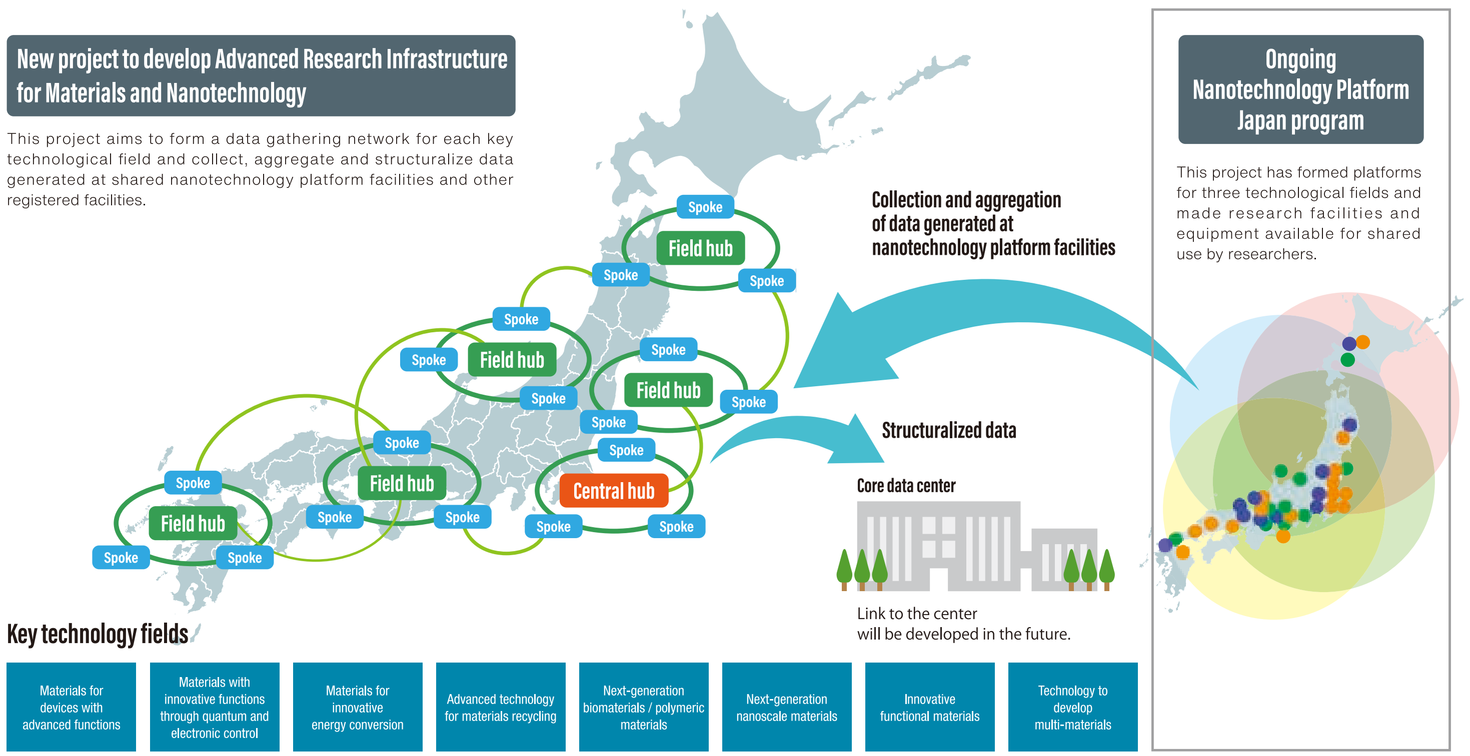
NIMS is highly qualified to lead this new project. Since its foundation 20 years ago, it has actively collected and used materials data, directed and supervised the Nanotechnology Platform Japan program and run its own advanced equipment sharing programs. Through these experiences, NIMS has learned that a wide range of high-quality data can be efficiently gathered by making its state-of-the-art equipment accessible to researchers and collecting their data. This lesson was effectively incorporated into the recently launched advanced materials research infrastructure project.

The project consists of a central hub, technology-field hubs and spoke functions. The spokes refer to research facilities at which high-quality data is actually generated. The field hub sites aggregate, format and structuralize high-quality data generated at the spokes and their own facilities. A group of spokes and field hubs is formed for each key technology field, allowing efficient generation of field-specific data using specialized equipment. NIMS, having assumed the role of the central hub, will work to optimize the operation of the entire project (see the diagram on p. 14).

For the project to succeed, data generation functions and the core data center will need to work together closely. "Data must be collected from advanced equipment by individuals who are knowledgeable about raw data handling, data structure, data formatting and ways in which data will be used," Koide said. "Top-down management is inappropriate for this project. It will be vital for the representatives of the central hub, field hubs and spoke functions to communicate closely. I will carefully listen to the opinions of data providers and users."

This project is scheduled to last 10 years. Advanced materials research infrastructure needs to be created in line with this time frame. "The data collection environment, including computing power and advanced research equipment, is expected to change significantly within the next 10 years," Koide said. "We need to create data collection infrastructure and generate data capable of withstanding these changes. In addition, it normally takes 20 to 30 years for the results of a research project to be put into practical use. A major goal of data-driven research is expediting this process. Ten years seem like a long time but it will pass very quickly. I hope that this project to create materials DX platforms will help Japan lead the world in the development of materials that can fulfill social needs."

(by Takeshi Komori)



Human involvement is the key to full-scale DX in materials development

The Advanced Research Infrastructure for Materials and Nanotechnology is indispensable for upgrading the materials research to be data-driven. According to Jun'ichi Sone, the Program Director for this infrastructure, human involvement is a critical component of the next-generation materials research infrastructure --a means of maintaining Japan's international competitiveness in materials R&D.

Jun'ichi Sone
Principal Fellow
Center for Research and Development Strategy
Japan Science and Technology Agency (JST)
Program Director,
Advanced Research Infrastructure
for Materials and Nanotechnology

Jun'ichi Sone graduated with a Master of Science degree from the University of Tokyo in 1975. He then joined NEC Corporation as a research staff of its Central Research Laboratories. He earned a Doctor of Science degree from the University of Tokyo in 1983. After serving as a General Manager of NEC's Fundamental and Environmental Research Laboratories, he became a Vice President of its Central Research Laboratories in 2007. He then joined NIMS as Executive Vice President in 2010 (currently Executive Vice President Emeritus of NIMS). He has been a Principal Fellow at the Center for Research and Development Strategy of JST since 2015. He is serving the Advanced Research Infrastructure for Materials and Nanotechnology as a Program Director from2021. He also served as a President of the Society of Nano Science and Technology Japan from 2012 to 2016. He won the Japan Society of Applied Physics Fellow Award in 2008.

Manufactured materials account for a large proportion of Japan's exports. Materials research is therefore important to Japan's industrial competitiveness and economic security. However, traditional materials research methods are becoming inadequate for that purpose, making the introduction of DX into materials research an urgent issue.

Popularizing data-driven research will require reviewing the ways in which shared research facilities are operated. These facilities need to be more accessible to researchers, and the data generated at these sites should be collected and aggregated more efficiently and processed into more user-friendly formats and structures.

Under the Nanotechnology Platform Japan program, we have supported Japanese materials researchers by making state-of-the-art equipment available to them. In the Advanced Research Infrastructure for Materials and Nanotechnology newly established, we hope to further upgrade the current cross-sectoral facility sharing systems by advancing DX in materials research across Japan. To achieve this, we have built data collection functions in the infrastructure consisting of hub and spoke structures in seven key technological fields absolutely vital to Japan's international competitiveness.

Materials design processes begin with exploration of materials which exhibits required function and performance and end with the development of optimum fabrication methods for the target materials. As part of materials design, first-principle calculations have become a common approach to determining the compositions and structures of materials. Next step of how to fabricate the derived materials is not easy because of enormous number of process methods and associated process parameters. The implementation of full-scale DX capable of materials exploration and fabrication process development is very challenging.

I believe human involvement will play a key role in overcoming these challenges. In line with the DX-related "humans-in-the-loop" concept, it is vital that human check and intervention mechanisms be integrated into digital data processing.

Experts skilled in the use of advanced equipment to optimize their operations who were trained in the Nanotechnology Platform Japan program will be deployed at the facilities of the Advanced Research Infrastructure for Materials and Nanotechnology. Incorporating their skills into materials DX platforms will be indispensable.

Streamlined data generation and utilization processes should greatly benefit the industrial sector as well. Material manufacturers may gain new insights into materials development by combining their proprietary data with data collected in this infrastructure. I am confident that strengthening coordination between the industrial and academic sectors based on this infrastructure will help the Japanese material industry maintain and improve its global competitiveness.

NIMS NEWS

English subtitles added to "To the Scientists of the Future" video clips

"To the Scientists of the Future" is a YouTube video clip series created jointly by NIMS and EUPHRATES Ltd. (a group specializing in creative work, including NHK's educational TV programs). We recently added an English subtitle option to these clips.

This scientific video series has been popular in Japan with a cumulative view count exceeding 7 million. These clips demonstrate the various intriguing scientific phenomena and unique materials NIMS has developed using fascinating images that are also entertaining and beutiful.

An English subtitle option makes it accessible to a broader global audience with an interest in science. We hope you enjoy it.

<How to display English subtitles>

1.

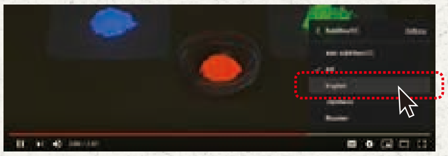
On the YouTube website, type "nims euphrates" in the search box and press enter. Select the video clip entitled "NIMS x EUPHRATES 未来の科学者たちへ."

2.

Left-click the "Settings" icon in the lower right corner of the screen and select "Subtitles".

3.

Finally, select "English" (only Japanese audio is available).



<List of "To the Scientists of the Future" video clips>

▼Sialon phosphor

▼Invisible glass

▼Super hydrophobic material

*Move video clips with English subtitles will be released as they become available.

Hello, my name is Laszlo, and I am from Hungary. I came to Japan in April 2017. I was a Postdoctoral Researcher in Kanazawa University before joining NIMS as an ICYS Fellow in February 2020. My research interest lies in green chemistry, in its broadest sense. Since the early period of my studies, I have been fascinated by the sophisticated processes that Nature developed to create materials with structural control at multiple length scales, starting at the molecular level. In ICYS, my research focuses on using biobased building blocks (polymers, small molecules) to create functional materials for high-end applications. It is the world-class research infrastructure and pioneering expertise in NIMS that prompted me to join this prestigious institute. In NIMS, I have unique opportunity to see my research from multiple perspectives, which greatly helps my scientific progress. With my research, I would like to contribute to the well-being of our society through environmentally more benign material concepts.

This is me enjoying my coffee with a spectacular view in Arashiyama, Kyoto.

Laszlo Szabo
(Hungary)
ICYS Research Fellow

NIMS NOW International 2021. Vol.19 No.5
National Institute for Materials Science

http://www.nims.go.jp/eng/publicity/nimsnow/

photo by Michihito Ishikawa(cover, p.2-3), Shinji Terakawa(p.4-7)
editorial design by Barbazio Inc.

To subscribe, contact:
Dr. Yasufumi Nakamichi, Publisher
Public Relations Office, NIMS
1-2-1 Sengen, Tsukuba, Ibaraki, 305-0047 JAPAN
Phone: +81-29-859-2026, Fax: +81-29-859-2017, Email: inquiry@nims.go.jp

R70 Percentage of Waste Paper pulp 70%
VEGETABLE OIL INK
ISSN 2436-3510