Promoting DOI Registration for Research Data
DOI Digital Object Identifier Subcommittee

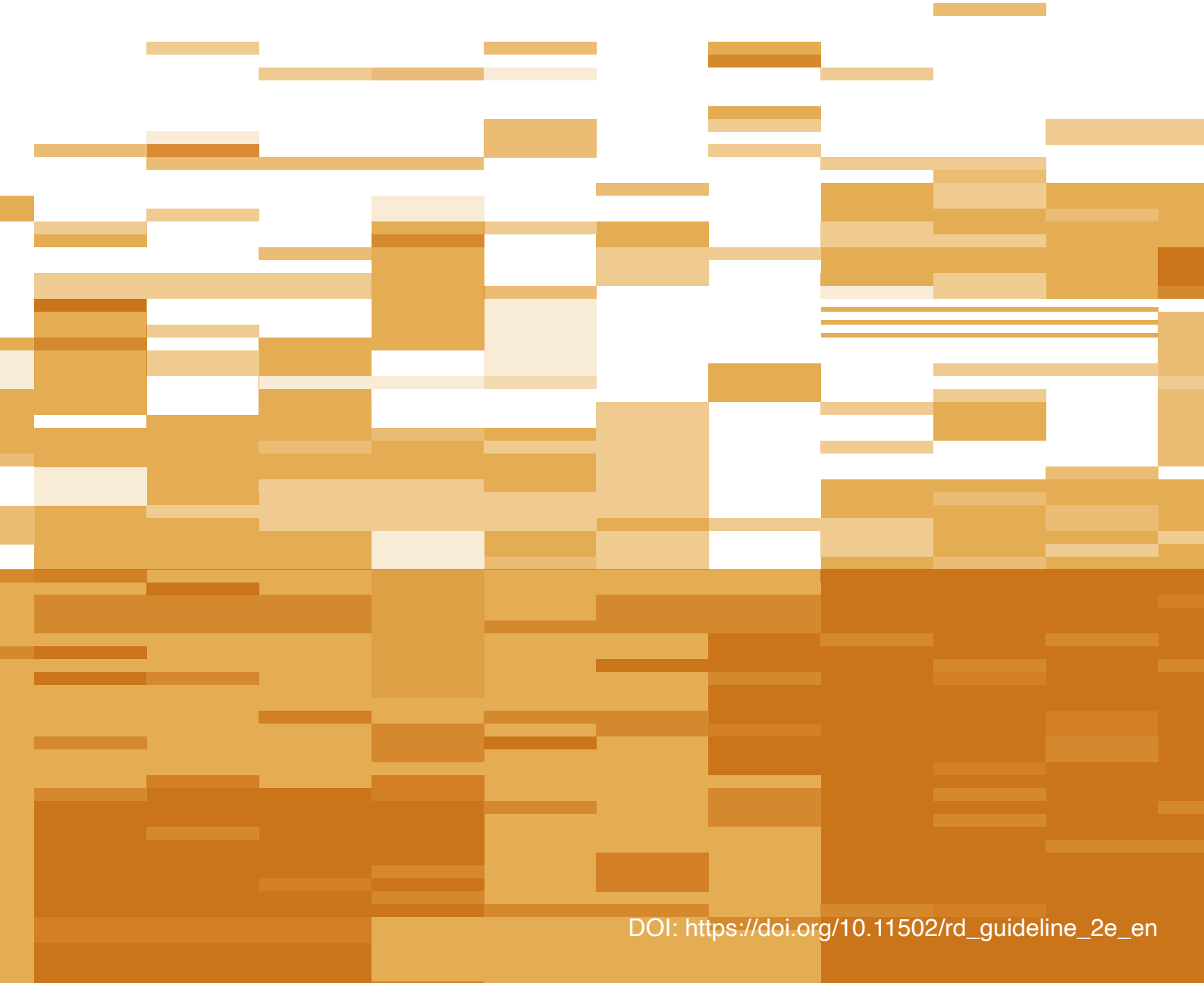
Japan Link Center

Guidelines for Registering DOIs for Research Data

June 3, 2024

Research Data Utilization Forum (RDUF)

Subcommittee for Promoting DOI Registration for Research Data



DOI: https://doi.org/10.11502/rd_guideline_2e_en

Revision History

Date of issue	Revision details	Issued by
October 20, 2015	First edition	Japan Link Center Steering Committee
June 3, 2024	Major revisions by the Subcommittee for Promoting DOI Registration for Research Data	Research Data Utilization Forum (RDUF) Subcommittee for Promoting DOI Registration for Research Data Japan Link Center Steering Committee

Contents

These guidelines are intended for interested parties who are considering registering DOIs for their research data through the Japan Link Center (JaLC). The main intended readers have been defined as follows (1) to (3) below; please use this as a guide for deciding which chapters to focus on reading.

For details on the specific registration operations of DOI and JaLC, please refer to Chapter 2 and the related documents listed in the references.

(1) Managers of organizations and institutions

The decision-makers for reviewing and determining DOI registrations on behalf of an organization and institution.

(2) Practitioners

Practitioners (repository operators, etc.) who actually register and manage DOIs and metadata.

(3) Researchers and data creators

Persons who wish to register a DOI on the data he/she is creating. Or the creator of the data for which the DOI is to be registered.

Table. Intended readers for each chapter

Chapter title	Main intended readers		
	(1) Managers of organizations and institutions	(2) Practitioners	(3) Researchers and data creators
1. Introduction	●	●	●
2. Overview of DOI and JaLC			
2.1 What is DOI?	●	●	●
2.2 DOI registration agency	●	●	
2.3 Overview of JaLC	●	●	
2.4 JaLC metadata		●	●
3. Policy on DOI registration for research data			
3.1 Significance and utilization of DOI registration	●	●	●
3.2 Data subject to DOI registration	●	●	●
3.3 Points to note when registering DOI		●	●
4. Procedures for registering DOI for research data			
4.1 Overview		●	●
4.2 DOI registration units		●	●
4.3 Creation and registration of metadata		●	●
4.4 DOI landing page		●	●
4.5 Operation after DOI registration		●	●

Chapter 1

Introduction

Persistent identifiers (PIDs), which uniquely and persistently identify objects, have played an important role in the distribution of information in the academic information field. Among all of them, the most representative PID is the Digital Object Identifier (DOI). Originally devised in the 1990s as a countermeasure against broken URL links, the use of DOIs has expanded with the development of the Web and electronic journals, and they have become indispensable for academic information, including identification of journal articles.

The DOI system is operated by the DOI Foundation, and DOI registration is conducted by DOI Registration Agencies (RA) accredited by the Foundation. In Japan, the Japan Link Center (JaLC), which was established with the aim of collecting academic content information originating in Japan and promoting its dissemination and utilization, was accredited as the only RA in Japan in March 2012.

Although journal articles are the main target of DOI registration, the scope is being expanded to include other research outputs (books, research data, etc.). At that time, various issues were being considered worldwide regarding DOI registration for research data, and Japan had little experience with this. JaLC conducted the Experimental Project on DOI Registration for Research Data from October 2014 to September 2015, inviting participation from the institutions involved with research data in Japan, with the aim of establishing a new system for registering DOIs for research data in Japan. This project identified and addressed issues specific to registering DOIs for research data, established operational methods, and considered ways to utilize DOIs. As a result, the first edition of these guidelines was created and published on the JaLC website in October 2015 [1].

Eight years have passed since then, and the open science movement, which aims to create new knowledge and accelerate and make research more transparent by promoting the publication

[1] Japan Link Center Experimental Project for DOI Registration for Research Data, "Guidelines for Registering DOIs for Research Data,"10/20/2015. [Online] Available: https://doi.org/10.11502/rd_guideline_ja.

and reuse of research data, has become more active internationally. In Japan, the Council for Science, Technology and Innovation has been playing a central role in discussions on the promotion of open science. In the field of research as well, an increasing number of academic journals are requiring the disclosure of supporting data when submitting journal articles, and in some cases, researchers are required to submit a data management plan (DMP) when starting research funded by external sources. This has forced researchers to seriously consider the management and disclosure of research data.

As the need for management, publication, and utilization of research data increases, the potential of DOI registration for research data is also gaining attention. The primary benefit of registering and publishing DOIs for research data is that it ensures persistent access to the data, increases interoperability, and promotes data usage. However, another benefit that cannot be ignored is that the use of DOIs makes it easier to track data citations and the number of data citations as well as the impact of the number of data publications and data utilization on the performance evaluation of data providers. As shown in Figure 1, the number of DOIs registered for research data is growing in Japan as well.

The revision of these guidelines were led by the Subcommittee for Promoting DOI Registration for Research Data, which was established in November 2021. As mentioned above, since the publication of the first edition of these guidelines, the registration of DOIs for research data has progressed in Japan. Therefore, in subcommittee meetings and open discussion forums, members with operational experience were asked to share their know-how and issues, while at the same time members who were trying to put the guidelines into practice for the first time were asked to raise questions and concerns, as well as shortcomings in the first edition of the guidelines. Discussions were held about the concerns of each participant, and the results were reflected in the new guidelines. The main revision includes the creation of a correspondence table for the chapters and their intended readers , a new description of the purpose of registering DOIs for research data, a description of “ways of thinking” for issues that do not have fixed solutions, such as “granularity,” and an expanded case study.

It is expected that the registration of DOIs for research data will become increasingly widespread in the future. We hope that the registration of DOIs for research data will promote the distribution of research data, contribute to increasing incentives for data provision, and serve as a driving force for open science as a whole. We hope that these revised guidelines will be of use for this in the field.

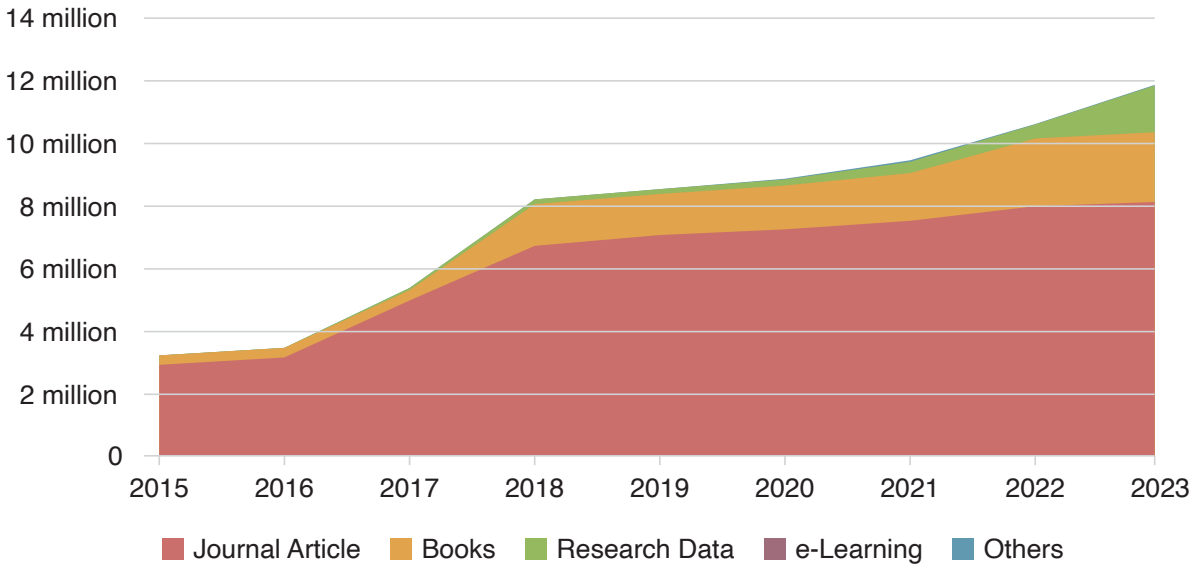


Figure 1. Accumulated number of the registration of DOIs for Research Data by JaLC (started in FY2015)

Chapter 2

Overview of DOI and JaLC

2.1 What is DOI?

The function of the DOI is to store a pair of an ID (DOI) assigned to a piece of content and its URL information, and to return the URL in response to a DOI query. If the URL of the content changes, the information in the pair is updated to ensure continuous access.

DOIs start with “10.” and are separated by “/”, e.g. “10.1241/xxx-oo-oo”. The “/” is preceded by a prefix, which is a unique string allocated by the DOI Registration Agency (RA) to the content manager. The part after the “/” is called as suffix, which can be determined by the registrant. Registered DOIs can be accessed as https://doi.org/“DOI” to resolve to the original URL and, as a result, to reach the content [2].

2.2 DOI registration agency

A three-tiered structure is adopted to ensure the sustainable operation of the system for DOI registration. The DOI system is governed by the DOI Foundation which has several organizations acting as DOI Registration Agencies (RAs) under it, and under those RAs are DOI registrants. JaLC is an RA, and its members are DOI registrants (institutions and organizations with no set duration of existence in principle, see 3.3.2 for details). (Figure 2-1) [3]

[2] The DOI Foundation, “DOI® HANDBOOK: DOI,” 4 2023. [Online]. Available: https://doi.org/10.1000/182.

[3] Japan Link Center Steering Committee, “What is the Japan Link Center? Its origins and basic policy,” 7/28/2014. [Online] Available: https://doi.org/10.11502/JaLC_policy.

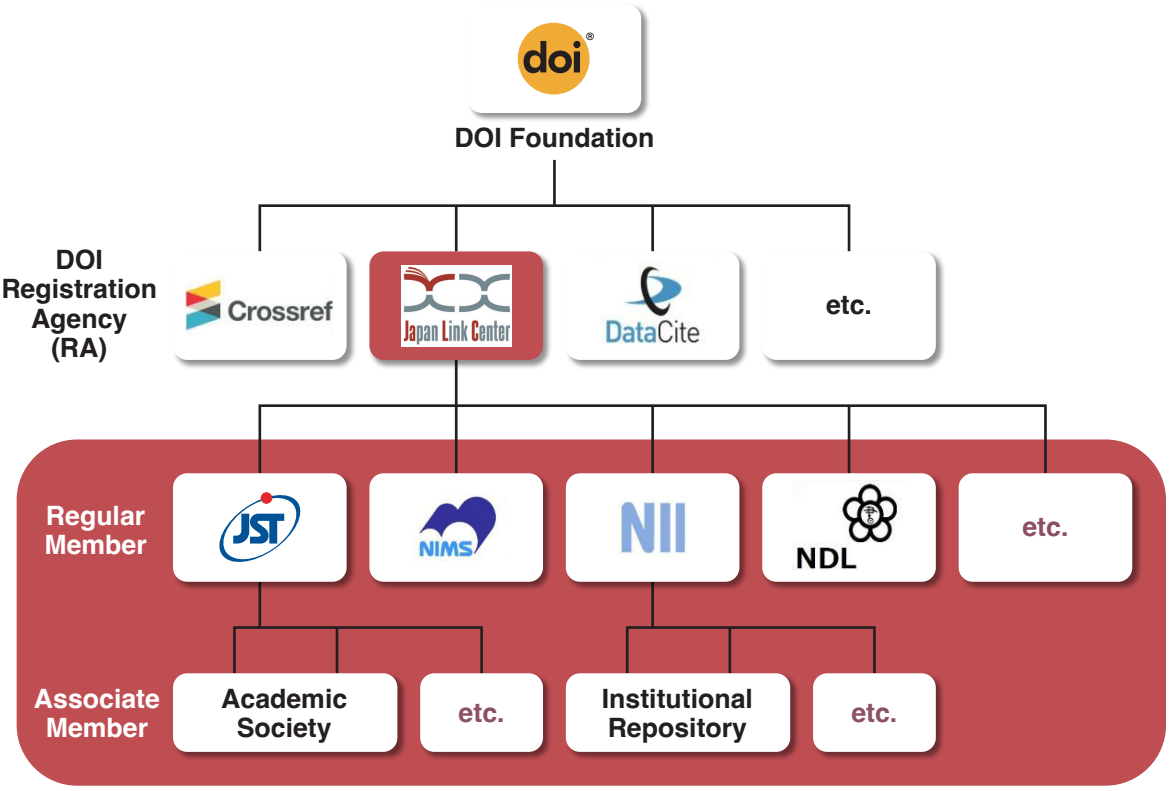


Figure 2-1. Organizational structure for operating DOIs

The DOI Foundation is responsible for managing a database of pairs of DOIs and URLs. Database management here means maintaining the functions of registering new DOIs and URLs, accepting changes to DOIs and URLs, and responding to DOI inquiries with URLs.

RAs, on the other hand, are authorized by the DOI Foundation to perform DOI registration work. As of March 2024, 12 organizations worldwide are registered as RAs. Each RA has its own registration policy regarding the type of content and scope of registrants and provides DOI registration services to the relevant registrants. In addition to JaLC, other major RAs include Crossref [4], which conducts large-scale DOI registration for journal articles, and DataCite [5], which was the first in the world in DOI registration for research data.

DOI registrants can register DOIs for their content by becoming members of individual RAs. DOI registrants are publishers or administrators of the content who pair the DOI with a URL to access the content and register it using a service provided by RA.

[4] Crossref, "You are Crossref: Crossref," n.d. [Online]. Available: https://www.crossref.org/.

[5] DataCite, "DataCite-Connecting Research, Advancing Knowledge," n.d. [Online]. Available: https://datacite.org/.

2.3 Overview of JaLC

JaLC aims to improve the convenience of domestic and overseas information services in Japan and promote the international distribution of domestic scholarly academic content produced in Japan or in Japanese by centrally managing bibliographic and location information on electronic scholarly academic content (journal articles, dissertations, books and reports, research data, e-learning, etc.) owned by domestic organizations and registering DOIs. As an RA, JaLC provides DOI registration services to JaLC members [6].

JaLC members are divided into two categories, “Regular Members” and “Associate Members,” each with different details of services available to them. Regular Members can register DOIs directly with JaLC, whereas Associate Members register DOIs with JaLC through Regular Members. JaLC assigns one or more unique prefixes to each Regular Member and Associate Member depending on the purpose (see 3.3.3) [7].

JaLC is an RA and a member of Crossref and DataCite. In addition to DOIs assigned by JaLC, JaLC members can also register DOIs assigned by Crossref and DataCite via JaLC (both require an additional fee). The flow is shown in Figure 2-2 [8].

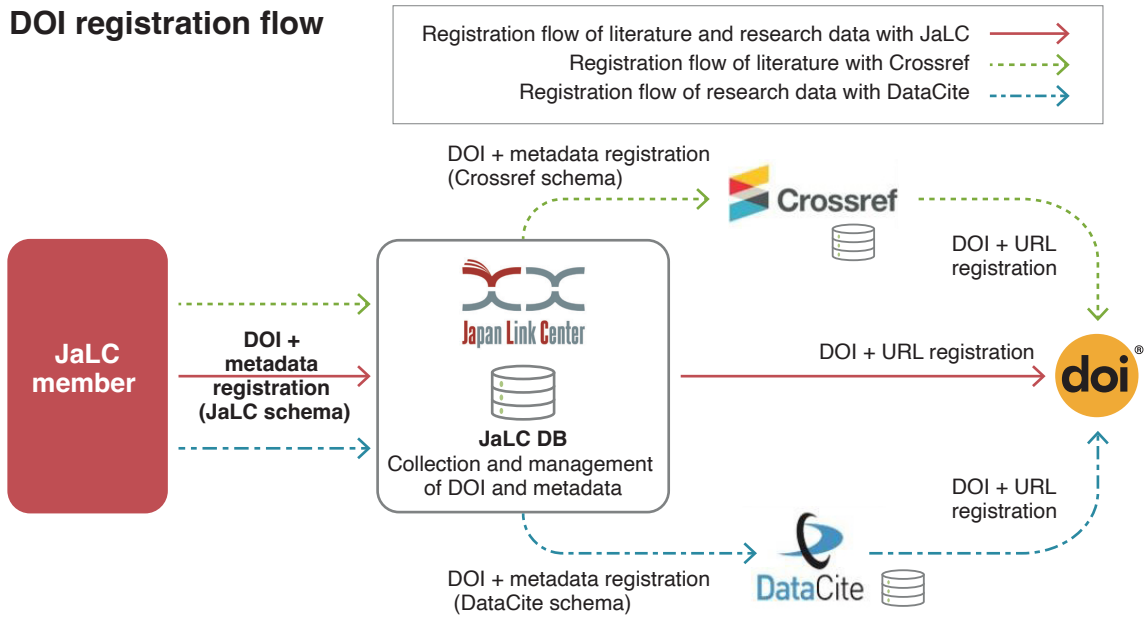


Figure 2-2. DOI registration flow via JaLC

[6] Japan Link Center, “Overview of JaLC: Japan Link Center (JaLC),” n.d. [Online]. Available: https://japanlinkcenter.org/top/about/index.html.

[7] Japan Link Center, “Membership type and annual membership fee: Japan Link Center (JaLC),” n.d. [Online]. Available: https://japanlinkcenter.org/top/admission/member_type.html.

[8] Japan Link Center, “DOI Registration: Japan Link Center (JaLC),” n.d. [Online]. Available: https://japanlinkcenter.org/top/service/service_doi.html.

2.4 JaLC metadata

JaLC defines metadata for each type of content, with different metadata schemas for journal articles, dissertations, books and reports, research data, e-learning, etc.

The metadata schema for research data corresponds to the DataCite schema (DataCite Metadata Schema 4.4) [9] (as of March 2024). DataCite aims to advance scientific research by providing researchers with a way to locate, identify, and cite research data, and it has defined a metadata schema to ensure that resources can be cited and retrieved accurately and consistently. Only some properties are mandatory, such as Identifier, Creator, Title, Publisher, Publication Year, etc., which are necessary for citation, to maintain compatibility with schemas in many fields. However, to make data more easily retrievable and citable, recommended and optional properties have been defined in addition to the mandatory properties. [10]

JaLC's metadata and metadata schema for research data are based on DataCite definitions, but to align the metadata and items with those of other content types (journal articles, etc.) that are subject to DOI registration, changes have been made to the DataCite definitions, such as how tags are named, how they are handled in multiple languages, and the addition of funding information. Additionally, items have been added to accommodate the common metadata items (as of March 31, 2023) defined in the "Basic Approach to the Management and Utilization of Publicly Funded Research Data" (Council for Promotion of Integrated Innovation Strategy, April 2021) [11].

*For information on creating metadata in JaLC,
see also **4.3.2 Creating metadata with distribution in mind.**

[9] Japan Link Center, "Japan Link Center (JaLC) Release Summary," 1/26/2022. [Online]. Available: https://japanlinkcenter.org/top/doc/ReleaseNotes_20220126_suppl.pdf.

[10] DataCite, "DataCite Metadata Schema 4.4," 30 5 2021. [Online]. Available: https://schema.datacite.org/meta/kernel-4.4/.

[11] Council for the Promotion of Integrated Innovation Strategy, Cabinet Office, Government of Japan, "Basic Approach to Management and Utilization of Publicly Funded Research Data," 2021. [Online]. Available: https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/sanko1.pdf.

Chapter 3

Policy on DOI registration for research data

3.1 Significance and utilization of DOI registration

3.1.1 Significance of DOI registration

The purpose of registering a DOI has both a conceptual significance and a practical purpose. First, conceptually, managing and storing data as academic objects and assigning them persistent identifiers (PIDs) called DOIs helps to preserve the uniqueness of the scientific discoveries represented by the data. In particular, registering a DOIs is a prerequisite for the publication of data in accordance with the FAIR principles, and by registering DOIs, persistent access to such data becomes more assured.

The FAIR principles include four principles for data management:
Findable, Accessible, Interoperable, and Reusable. [12]

- **Findable:**
  Ensure that the data can be easily found. It must be clear where the data is, what it relates to, and how it can be accessed. It is also important that metadata and identifiers are in place to allow data to be searched.
- **Accessible:**
  Data should be managed to be as widely accessible as possible. Ensure that as many people as possible can access and utilize the data.
- **Interoperable:**
  Data should be interoperable across different systems and applications. Data should preferably be shared (internationally or within a community) in a standardized format and managed so that it is compatible across different programs and platforms.

[12] Junichi Onami, Shigeru Yatsuzuka, Tomoe Nobusada, Mari Minowa, Nobutaka Mitsuhashi, Hideki Hatanaka, "The FAIR Principles as a Standard for Data Sharing," Bioscience Database Center, 2018. [Online]. Available: https://doi.org/10.18908/a.2018041901.

- **Reusable:**

  Data should be reusable. The data must be managed in such a way that those who obtain it can reuse it and use it repeatedly thereafter.

The FAIR Principles provide guidance to promote best practices in data management, and the DOI as a permanent identifier is the foundation that supports the FAIR Principles.

On the other hand, the practical objective is to have the data actually found and used, cited and evaluated, and registration of DOIs is considered the most efficient method for this purpose. Registering DOIs makes it easier for researchers, academic publishers, data repositories, etc. to access and cite research data. DOIs can also be used to track the utilization of research data. This allows to measure and evaluate the impact of research data.

In addition, DOIs can help improve the reliability and quality of research data. Research data with DOIs are more reliable in that they are semi-persistently accessible. On the other hand, it is expected that research data with DOIs are subject to adequate quality control to ensure the persistence and reliability of such data.

3.1.2 How to utilize DOI registered for research data

The following methods are available for utilizing DOIs registered for research data.

- **Data citation:**

  Researchers and academic publishers citing DOIs for research data may serve as a reference source for studies by other researchers. This is expected to increase the reliability of the research and increase its impact.
- **Data retrieval:**

  The use of DOIs for research data facilitates retrieval and access to the data. Many data repositories and publishers use DOIs, allowing access to data simply by entering the DOI.
- **Data management:**

  Using DOIs for research data allows researchers to easily manage their data. Data with DOIs are guaranteed to be persistently accessible, which facilitates management to ensure data security and persistence.
- **Data evaluation:**

  DOIs for research data can be used to track utilization of the data. This allows to measure and evaluate the impact and influence of research data. Data with a DOI facilitates evaluation of the data because it is easy to track utilization of such data.
- **Data sharing:**

  DOIs for research data facilitate data sharing. Data with DOIs are guaranteed to be persistently accessible, allowing researchers and data users to share and reuse research data.

3.2 Data subject to DOI registration

3.2.1 Concept of data subject to DOI registration

DOIs provide assurance of the accessibility and existence of data.

It is desirable for experts and researchers to select data that is deemed important for the promotion and development of science and academics in the relevant field, and that requires long-term preservation and persistent access. There is a proposal to use peer review to decide which data should be assigned a DOI and made persistently accessible, but at present the policy, methodology, etc., are still far from finalized.

It will be necessary to increase the number of cases examined individually by communities, organizations, and relevant stakeholders.

In addition, an increasing number of journals are requiring as part of their publication policy that data be accompanied by a DOI upon submission and publication of manuscripts. Such data should be handled in accordance with the editorial policies of each publisher and journal. In recent years, there have been cases where data are registered and published as DOIs together with editorials and commentaries on the research data themselves in the form of “data journals”. Similarly, research funding agencies may require publication of resulting research data and registration of DOIs.

3.2.2 Concept of quality

DOI registration does not generally provide any assurance of the quality of the data itself, although information about the quality of the data may be included in the metadata. In addition, in terms of guaranteeing uniqueness, the DOI registration itself may contribute to quality control.

In some cases, the data to be registered with a DOI is selected through peer review by experts. However, there is no uniform approach to quality control for all data subject to DOI registration.

As mentioned in Chapter 3.2.1, for data candidates for DOI registration, consideration should be given to ensuring the quality of meaningful data in accordance with the FAIR data principles. Checking and managing data quality will likely require the involvement of researchers and experts with expertise in the field. It is desirable for data generators, managers, providers, etc. to ensure this. In field-specific data repositories, it is recommended to conduct data quality checks by experts in relevant field (researchers, data curators, etc.) as part of the curation process. Comprehensive data repositories that are not restricted to a specific field also exist and are utilized, but it is important to note that these repositories do not perform data quality evaluation.

3.2.3 Approach to coordinating DOI across multiple data repositories

It is not recommended to designate multiple DOIs for the same target data. In principle, only one DOI should be registered for each object. However, the DOI system allows multiple DOIs

to designate the same data or object [2]. In the case of research data, this may also need to be considered in special circumstances. For example, multiple DOIs may be registered for the same research data due to circumstances such as the involvement of multiple institutions in data creation or the registration of a new DOI due to the operational rules of the data repository where the data are registered. In such cases, it is desirable for data providers to communicate with each other and understand each other's situation, to keep track of each other's DOI registration status, and include such information in the metadata. It is recommended to state in the relation information of the metadata for the related DOI (the relation list field of the metadata item) that another DOI has been registered for the same research data.[13]

3.3 Points to note when registering DOI

3.3.1 Subjects of DOI registration and data management

When registering a DOI for research data, it is necessary to manage both the DOI registration and the research data being registered. Keep in mind that the respective administrators may be different in an organization.

The manager of the DOI registration (hereinafter referred to as “DOI manager”) is responsible for creating and maintaining metadata and landing pages (including updating them when necessary), maintaining access to content URLs associated with DOIs, and other management necessary for the DOI system to operate properly (JaLC Terms of Participation, Chapter 7 "General Obligations of Regular Members") [14]. Although there may be multiple DOI managers within an organization, the organization as a whole is centrally managed by the person (department) registered with JaLC. On the other hand, the content itself (in this case, research data) is properly managed by a manager designated within the organization. DOI managers and content managers are encouraged to fully share information and cooperate to manage the data.

This section 3.3 describes the matters to be taken care of by the DOI manager and discuss points that require attention, focusing mainly on the differences from registering DOIs for journal articles. The management of content (research data) is not covered in the scope of these guidelines; only reference information [11] [15] is provided at the end. As stated in Section 3.1, it is recommended that management be conducted in accordance with the FAIR principles.

[2] The DOI Foundation, "DOI® HANDBOOK: DOI," 4 2023. [Online]. Available: https://doi.org/10.1000/182.

[13] Japan Link Center, "JaLC Content Registration Manual Web Interface Version [Research Data]," 2023. [Online]. Available: https://doi.org/10.11502/web_interface_manual_research_data.

[14] Japan Link Center Steering Committee, "Japan Link Center Terms of Participation," 2023. [Online] Available: https://japanlinkcenter.org/top/doc/jalc_sankakiyaku.pdf.

[11] Council for the Promotion of Integrated Innovation Strategy, Cabinet Office, Government of Japan, "Basic Approach to Management and Utilization of Publicly Funded Research Data," 2021. [Online]. Available: https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/sanko1.pdf.

[15] Cabinet Office, Japan Open Science Promotion Committee, "Guidelines for Development of Data Policies for National Research Institutes," 6/29/2018. [Online]. Available: https://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.html.

3.3.2 DOI management system

Membership in the JaLC, which is required for DOI registration, is available to organizations that, in principle, have no set duration of existence, due to the need to ensure the persistence of access to content. Member organizations shall appoint a responsible department to serve as a contact point and coordinator for DOI registration.

When the registered content is an article, the content is generally owned by a publisher, university, academic association, or other organization with no set duration of existence. However, in the case of research data, the target content may be generated from a fixed-term project, and the direct manager may not necessarily be an organization with an indefinite term of existence. In that case, it is necessary to consider a system for guaranteeing access to data with registered DOIs over the long term. For example, the following measures are possible.

- The participating institutions of the project (not the project itself) become Regular or Associate Members of JaLC and register their DOIs with JaLC as JaLC members .
- If the project is long term, consider making the project itself a JaLC member. Upon application, the JaLC Steering Committee may approve a project without an organization to become a JaLC member after reviewing its sustainability and other factors.
- When a project registering DOIs for research data involves multiple institutions, the institution representing the project is responsible for managing DOIs. In this case, the DOI registration process may be carried out either by the central institution collecting all the data or by each institution individually, but in any case, the registrant should be clearly assigned within the project. The institution registering the DOI must be entitled to be a Regular or Associate Member of JaLC. The project representative must know which research data from which participating institutions will have DOIs registered.

Below are examples of DOI management systems for single and multi-institutional projects.

- **Examples of projects within a single institution**

  Figure 3-1 shows an example of a project within a single institution. Project P within Institute A is a case in which new DOIs are registered for the research data obtained through the project. Institute A has registered DOIs for journal articles and is already a JaLC member, it will register DOIs for the research data generated by Project P based on that qualification. At this time, the organization applies to JaLC for a prefix for research data separate from the prefix assigned for registering DOIs for journal articles, and the DOI is registered using this prefix (prefixes are explained in the next section “3.3.3 Prefix assignment policy”). It would be easier to operate separate management systems for journal articles and research data, given that they are often considered to be managed separately, even within the same institution. After the project is completed, Institute A will be responsible for ensuring access to the data as an institution.

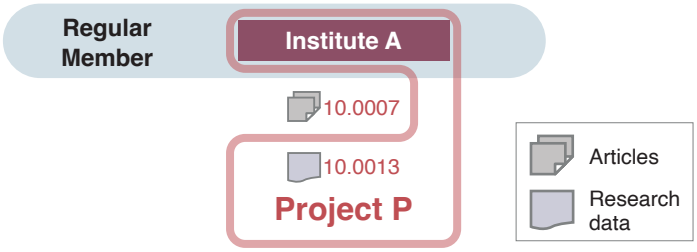


Figure 3-1. Example of a project within a single institution

- **Example of a multi-institution project**

Figure 3-2 shows an example of a multi-institution project Q. The project representative organization is Institute A, which is a Regular Member of JaLC. In addition, the other institutions participating in the project, Society X, University Y, and University Z, are Associate Members of JaLC under the Regular Members of JaLC (Society X under Institute B, universities Y and Z under Institute C), which are different from Institute A. If Institute A collects all research data generated by Project Q and registers a DOIs, then only Institute A will apply for the prefix to JaLC and will state in the application that it is the representative institution for the project. Alternatively, if it is decided in a project that Institute A will not coordinate the registration of DOIs, but that each institution will do so on its own, then the Associate Member Society X will register their DOIs through the Regular Member Institution B, and the Associate Members universities Y and Z will register their DOIs through the Regular Member Institute C. Regular Members apply for a prefix with the project information to the JaLC Secretariat, and then register the DOI.

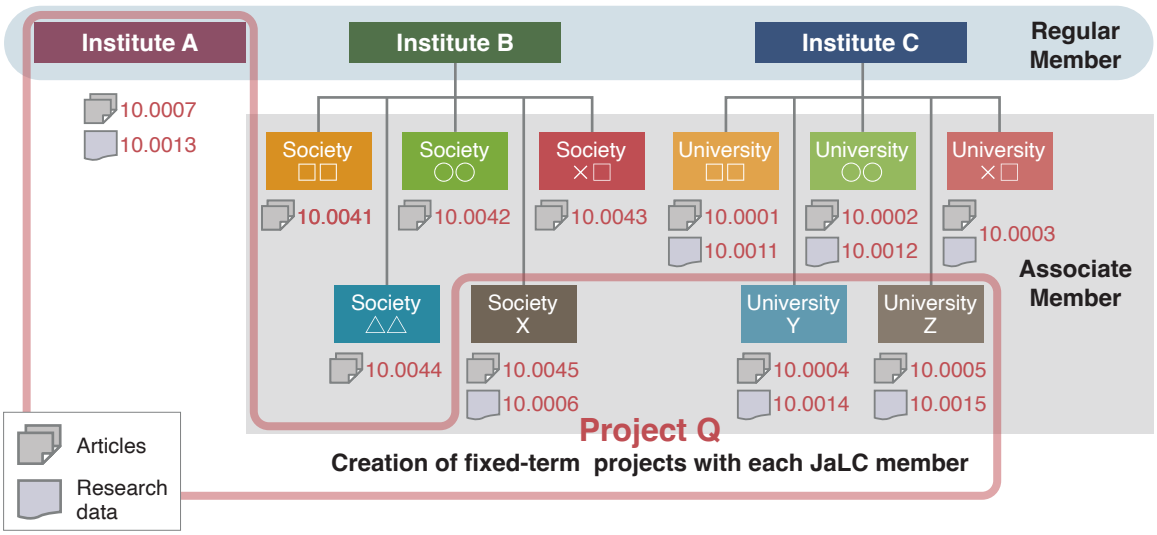


Figure 3-2. Example of a multi-institution project

The above are sample cases for the purpose of explanation. In practice, it is desirable to design DOI names, including prefix applications, and develop operational procedures in accordance with the circumstances of each institution and project, using these examples as reference.

3.3.3 Prefix assignment policy

DOI prefixes are issued by the Secretariat to JaLC Regular Members. As a general rule, one prefix is issued per member, but multiple prefixes can be issued if necessary.

The creation and management of research data differs from journal articles and other literature in that a variety of cases can be anticipated, such as the data creator being a single laboratory within an institution or the participating institutions being a fixed-term project consisting of multiple institutions. Therefore, considering the possibility that data management may be transferred to another institution in the future, it is considered easier to operate the system if separate prefixes are used for each administrator of the actual data or metadata. For example, in the case of a multi-institution project, the prefixes may be used for each of the participating institutions; in the case of a joint repository, the prefixes may be used for each of the institutions creating research data, as necessary.

Also, different prefixes may be applied for different content types. For example, if an organization has already registered a DOI for a journal article as a JaLC member, but it is registering a new DOI for research data and has a separate administrative department, it would be easier to manage the DOIs by using a different prefix.

If a new prefix is required, an application must be submitted to the JaLC Secretariat along with details of the registration subject, etc., and a new prefix will be assigned. Members must ensure proper management of access and metadata for data for which they have registered DOIs. In addition, Regular Members must report the status of prefix management to JaLC once a year and return any prefixes that are not in use and are not expected to be used in the future.

3.3.4 Guarantee of persistence of access

As described in “3.3.2 DOI management system”, the DOI registration is performed by JaLC members among the institutions participating in a project. Such members must prepare measures to ensure continued access.

In some cases, research data is managed on a small scale, such as at a research laboratory, and special measures must be taken to ensure long-term access. In the case of collaborative projects with other institutions that are carried out outside the institution's organizational structure, the issue of which institution will manage the data after the project has ended may arise. If the DOI manager or content manager cannot guarantee sustained access, it may be possible to ensure access by outsourcing data management to an appropriate relevant institution or data center.

(Examples of data storage after project completion)

- Example 1: The representative institution takes over and manages all project data with registered DOIs. This may involve a transfer of prefixes. In addition to the case that researchers who registered DOIs store the data, institutional repositories, libraries, etc. may take that role.

- Example 2: The institution that registered the DOI stores the data.
- Example 3: If there is no one to accept the data, the data may be deposited in an external data repository, etc. It will also be effective to establish a system that will enable continuous access during and after the project is completed through regulations.

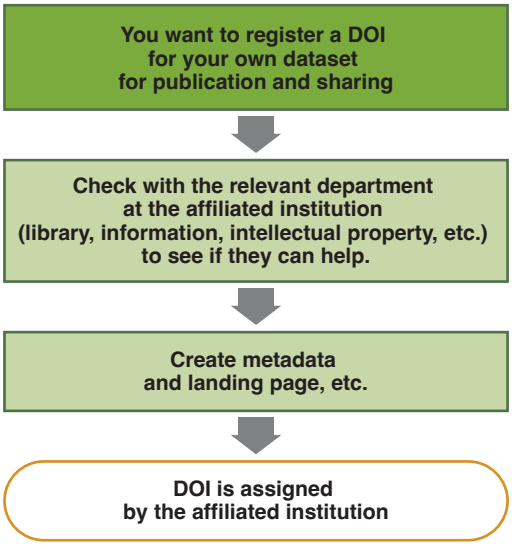
Chapter 4

Procedures for registering DOI for research data

4.1 Overview

This section describes the actual procedure for registering DOIs for research data from the perspective of the researcher who owns the research data. Figure 4-1 shows a schematic diagram of the process of registering data and metadata in an affiliated institution's repository and registering a DOI. If your institution does not have a data repository, see the separate article “How to assign DOIs to research data?" [16].

Refer to "5. Case Studies" for examples of actual registration procedures flow for different institutions.



(Research Data Utilization Forum（RDUF）, Research Data Citation Subcommittee. Modified from “How to assign DOIs to research data?”: https://japanlinkcenter.org/rduf/doc/rd uf_rdc_doileaflet.pdf) [16]

Figure 4-1.
Flow of DOI registration for research data at the affiliated institution (diagram)

[16] Masahito Nose, Kazuki Ohmukai, Mizuho Owashi, Nanako Takahashi, & Yasuhiro Murayama, "How to assign DOIs to research data? A five minute guide to assigning DOIs to research data," Research Data Citation Subcommittee of the Research Data Utilization Forum (RDUF), 2020. [Online]. Available: https://doi.org/10.11502/rduf_rdc_doileaflet.

If a researcher wishes to register a DOI for their own dataset for publication and sharing, they must check with the relevant department at their affiliated institution to see if this is possible, taking into consideration external factors such as journal article submission guidelines and project rules, as well as their affiliated institution's data policy and repository maintenance status. Since multiple departments may be involved in DOI registration, such as the library and the information/ intellectual property department, it may be necessary to individually confirm which department is responsible for which task. After the relevant department has finished reviewing the data, the researcher then performs such tasks as creating metadata and landing pages, etc. Some institutions may be able to provide assistance with creating metadata and landing pages. The following sections explain the points that require special attention assuming such cooperation with multiple departments.

4.2 DOI registration units

4.2.1 Basic concept

Metadata and landing pages correspond one-to-one to one DOI, but the actual data to be registered may consist of multiple files. When registering DOIs for journal articles, it is possible to handle them in easy-to-understand units, with one DOI registered for each journal article. However, in the case of research data, the nature and usage of data varies widely, so it is difficult to determine at what granularity DOIs should be registered. Ultimately, the granularity is at the discretion of the data provider. In doing so, it is necessary to remember that DOIs will be used for a long time.

4.2.2 How to determine granularity

Points to consider when determining the granularity of DOI registration are listed below.

- **Citation:**

  Linking citations and counting citations are one of the primary purposes of DOIs. Registering DOIs for research data will encourage data citation and make it easier to measure results. Therefore, it is important to register DOIs at the granularity at which data providers expect to be cited. When DOIs are registered for data units that are often used together, it will be easier for data users to cite them.
- **Nature of data:**

  It is desirable that the granularity should be meaningful on its own. However, the granularity of what is meaningful on its own may vary depending on the type of data, such as observational, experimental, or computational data. For example, most observational data are difficult to reproduce, and a single measurement does not necessarily have meaning on its own, but experimental data are expected to be reproducible under the same conditions and may have

meaning on a per-experiment basis.

- **Ease of access:**

  Aside from fulfilling the role of the DOI in ensuring access to content, it is desirable to register DOIs at a granularity that is easily accessible to users in order to ensure that content is effectively referenced and used.

- **Ease of management:**

  It is also important to ensure ease of management so that continuous access can be guaranteed. Consideration should be given to making the data easy to be managed by data managers and institutions which provide the research data. Care must be taken to ensure that no data are left without a manager even though the size of research data may be large and may have been measured and created by multiple institutions.

- **DOI quantity:**

  The number of DOIs should be kept at a level that allows the DOI registration and resolution system to operate properly. Note also that the annual membership fees for the DOI Foundation and DOI registrars will increase with the number of registrations. Therefore, if the number of DOI registrations will be huge (hundreds of thousands per year), it is advisable to consult JaLC in advance.

4.2.3 How to assign DOI suffixes

The suffix is simply a string of characters used to uniquely identify a DOI and has no inherent meaning or structure. Managers may use any structure for their convenience, so in some cases randomly generated strings may be used, but in other cases the strings may also be used to indicate the meaning and structure of data, making them easier for humans to understand and facilitating data management and data utilization. Because DOIs are a long-term means of access, careful consideration should be given when including names that are expected to change (such as organizational department names or project names).

- (prefix)/(random number)
  Example: 10.17591/55838dbd6c0ad
- (prefix)/(name of sub organization).(DB name) - (identifier in DB)
  Example: 10.14977/05.tdbs-23732
  10.14977/05.gsj-aster-xxxx
- (prefix)/(institution abbreviation).(Serial number)
  Example: 10.11503/nims.1001

There are also examples where the suffix reflects the relationship between datasets, as shown in the following chapter.

- (prefix)/(repository name). (dataset ID)-(entire dataset (000) or data subset number).(version number).

Example (entire dataset): 10.18908/lsdba.nbdc01530-02-000.V010

Example (data subset): 10.18908/lsdba.nbdc01530-02-001.V010

In this case, if a DOI is registered for a collection of multiple datasets, a DOI can also be registered for each subset, making it possible to cite the entire dataset as a collection as well as each subset.

4.2.4 If you want to show the relationship between multiple data

Although DOIs as identifiers are independent of each other, there is an emerging need to show the relationship between multiple datasets for the convenience of users and for data citation. Examples of operational ideas are shown below for your reference.

- **Primary and secondary data:**

  For example, when secondary data generated from multiple primary data are published, especially when the primary data are provided by multiple institutions, there is a need for a mechanism to evaluate the indirect use of the primary data when the secondary data are cited. To address this issue, a method has been proposed in which DOIs are registered for primary data and their relationships are described in the metadata item “Relation” [17], and this method is beginning to be used in practice [18]. It is recommended to refer to the JaLC Content Registration Manual and select the appropriate option for the Relation item according to your purpose [13]. In this case, “IsDerivedFrom” can be used.

- **Full set and subsets of an aggregate dataset:**

  When a DOI is registered for a collection of multiple datasets, it may be required to cite only a subset of the entire dataset as a collection that has actually been used. One approach to this is to register DOIs for both the full set and the subsets and indicate the relationship between the datasets through suffixes and landing page descriptions (see the last example in the previous chapter on suffixes). The relationship can also be expressed using the option "IsPartOf" of the metadata item "Relation."

In addition to this, there is also a case of having two DOIs registered at the time of data registration: a DOI specific to that data (Concept DOI) and a DOI used for version control (Version DOI), so that multiple versions of the data can be cited separately [19].

[17] Yasuhiro Murayama, Masahito Nose, Yukinobu Koyama, Hideaki Takeda, Koji Zettsu, Takenari Kinoshita, Toshihiko Iyemori, and Takashi Watanabe, "Proposal of a new DOI citation method for dataset generated from other datasets," 138th SGEPSS General Meeting and Lecture, Society of Geomagnetism and Earth, Planetary and Space Sciences, 2015.

[18] S.Murakami, S., M. Yamada, A. Yamazaki, K. McGouldrick, Y. Yamamoto, G.L. Hashimoto, "Venus Climate Orbiter Akatsuki UVI Calibrated Data PDS3 dataset," 11 11 2017. [Online]. Available: https://doi.org/10.17597/ISAS.DARTS/VCO-00003.

[13] Japan Link Center, "JaLC Content Registration Manual Web Interface Version [Research Data]," 2023. [Online]. Available: https://doi.org/10.11502/web_interface_manual_research_data.

[19] Zenodo, "Frequently Asked Questions: Versioning," n.d. [Online]. Available: https://help.zenodo.org/faq/.

4.3 Creation and registration of metadata

4.3.1 Basic concept

Metadata refers to "data that describes data" and is important additional information for managing and utilizing data. Identifiers such as DOIs are used for distribution purposes, but without metadata that explains and describes what the target object is, classification and retrieval would be difficult, so metadata plays a significant role.

Metadata for research data is often created with data utilization in mind. The metadata associated with the research data is classified into the following categories.

- Metadata created as necessary when research data are generated, analyzed, used, and managed in the course of research (“research metadata”).
- Metadata created when research data are shared and published in repositories, etc. (“publication metadata”).
- Metadata created to facilitate data distribution (“distribution metadata”).

As shown in Figure 4-2, there is a lot of overlap in the metadata, but necessary items may vary depending on the purpose.

The next chapter discusses the creation of the distribution metadata, which is essential when registering DOIs for research data.

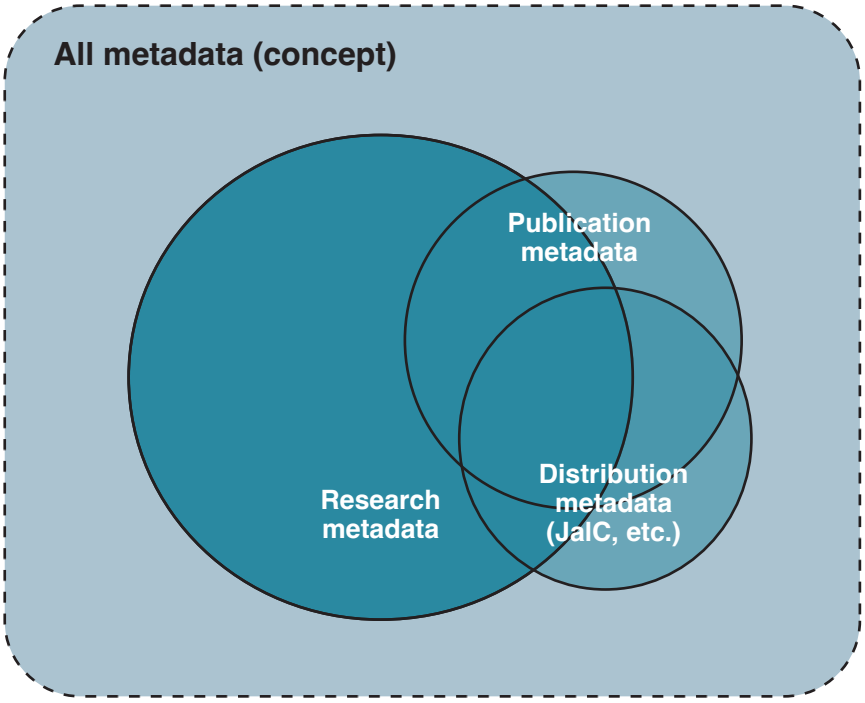


Figure 4-2. Relationships among metadata associated with the research data

4.3.2 Creating metadata for data distribution

First, the initial purpose of registering a DOI and publishing data is to distribute the published data. In addition to collecting and managing the content, it is necessary to collect and organize necessary information in anticipation of the distribution of data in accordance with the FAIR principles.

The metadata schema provided by RA for DOI registration is also defined with data distribution in mind. Therefore, if the managed information can be properly described and input according to the schema, you can enjoy various benefits in distributing academic information for which DOIs are registered.

When obtaining a DOI through JaLC, the information must be entered according to the JaLC schema. Note that the required items differ depending on the type of subject being registered (journal articles, research data, etc.). Specific metadata items are described in detail in the following manuals, which we encourage you to use.

JaLC Content Registration Manual Web Interface Version [Research Data] (in Japanese) [13]
JaLC Content Registration Manual XML Format Guide [Research Data] (in Japanese) [20].

*See also 2.4 "JaLC metadata" for the metadata of JaLC.

4.3.3 How to make it easier to enter information

Creating metadata requires a great deal of effort, so it is desirable to reduce the labor required for creation by effectively selecting metadata items and using tools, etc.

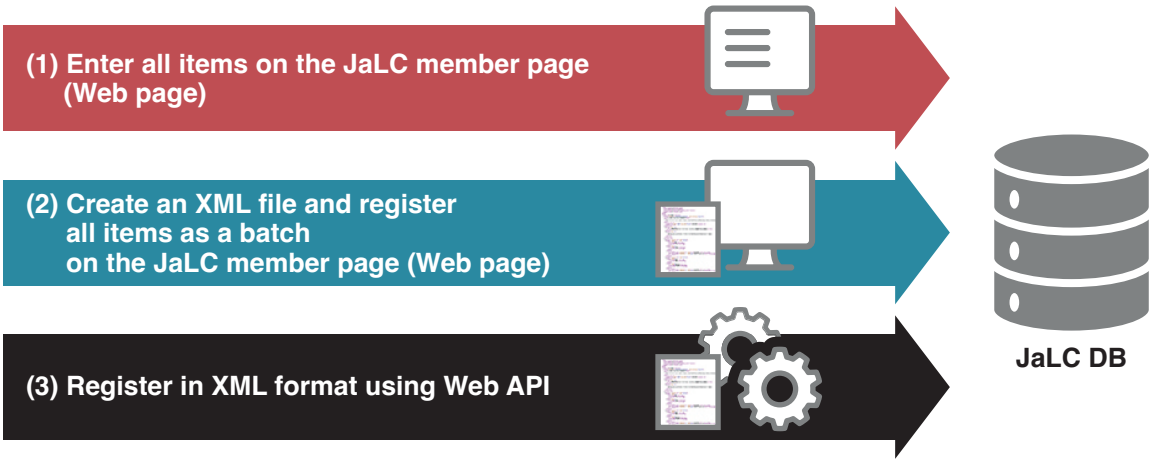


Figure 4-3. DOI registration method at JaLC(Source: JaLC website [8])

[13] Japan Link Center, "JaLC Content Registration Manual Web Interface Version [Research Data]," 2023. [Online]. Available: https://doi.org/10.11502/web_interface_manual_research_data.

[20] Japan Link Center, "JaLC Content Registration Manual XML Format Guide [Research Data]," 2023. [Online]. Available: https://japanlinkcenter.org/manual/research_data/index.html.

JaLC has a web interface for DOI registration as described in the manual above, which allows interactive and easy-to-understand information entry (1).

Other methods include creating XML files for registration and verification (2) and batch registration using an API (3), so a more efficient method can be used depending on the volume and nature of the data.

Additionally, DOI registration practitioners can provide assistance to researchers and data creators, including, for example, the following. Refer to section 5 "Case Studies" for concrete examples.

[Providing your own online form]

- Prepare an easy-to-understand online form specifically for information entry and have researchers and data creators use it to enter metadata.
- Use selective vocabulary when entering data, so that it is widely known in the community.
- Create an input system that is linked to the input GUI and has the function to export metadata according to the schema to automate the creation of metadata.
- Enter metadata etc. into a web UI form and create it while checking where it will be output in the metadata format according to the schema.

[Introduction of batch system]

Although data entry using the online form is done one item at a time, if you are registering DOIs for a large amount of existing content at once, it may be effective to arrange the required metadata items as columns in a format such as CSV, and process them in batches, with one record per line.

4.3.4 Notes

While the purpose and benefits are clear, there are often various limitations in creating metadata, including difficulties in defining common items, restrictions regarding disclosure, and difficulties in obtaining information when the data creator and metadata creator differ. To address these limitations, it is recommended that researchers and data creators indicate best practices or guidelines and follow them as much as possible when creating metadata.

In many cases, such as when data is published as institutional policy, the data manager, not the data creator, takes the initiative in creating and managing metadata, so ensuring governance is also an important aspect.

4.4 DOI landing page

When a registered DOI is accessed, the Web page to which access is redirected is called the landing page of such DOI. A corresponding landing page must exist for each registered DOI. The data themselves for which the DOI is registered may be either public or restricted access, and this distinction between open or restricted access should also be noted on the landing page. In addition, even though the data themselves is no longer available, the landing page should continue to be open to the public to ensure persistent access for the registered DOIs. See also section 4.5.5 for withdrawal of data publication.

In addition to the five mandatory properties (DOI, Creator, Title, Publisher, Publication Year) required by the JaLC metadata schema, the following are recommended as information to be included on the landing page.

- Summary description (abstract) of the data
- Contributor(s) who contributed to the content of the resource
- Information on how to cite the data
- Public or restricted access
- History of additions, modifications, updates, withdrawals, divisions, mergers, etc.
- Type and format of the data
- License information for data utilization (Creative Commons, displaying under the institution's own license system, etc.)
- Relations to other related resources, such as citing/cited, including/included, etc. (Relation)
- Other field-specific metadata information

In addition to providing metadata for circulation, it is also possible to include some of the metadata information (research metadata) that will be useful to researchers and experts in the same field as well as related and different fields, which is expected to improve the usefulness of research data (see Figure 4-2).

As explained in Chapter 4.2, for example, if the a set of research data that has different versions, and a DOI is registered for each version, a landing page must be provided for each DOI. However, it is possible to choose whether to provide a separate landing page for each DOI on a one-to-one basis, or to share the landing pages of related DOIs on an N-to-one basis.

There may be several ways to generate landing pages, including the static method of creating and publishing individual HTML files, the dynamic method of using a dedicated program to generate landing pages from a research data database management system, or the use of a content management system (CMS). It is advisable to carefully select an implementation method that best suits the use case.

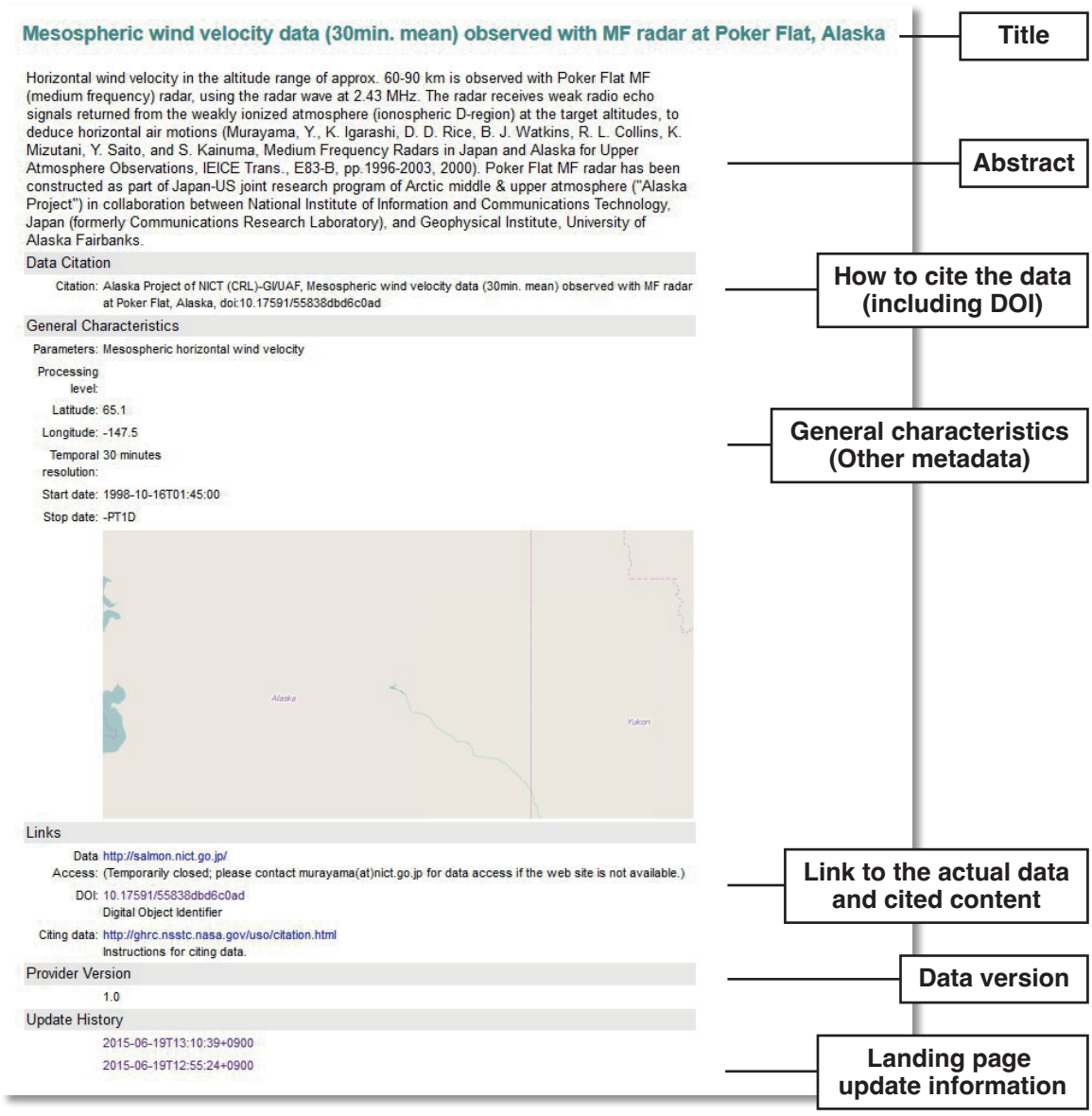


Figure 4-4. Example of a landing page (Source: doi:10.17591/55838dbd6c0ad)

4.5 Operation after DOI registration

4.5.1 ORCID integration

Data for which a DOI has been registered through JaLC can be registered as performance information for data creators with ORCID, a service that provides researcher identifiers as performance information for data creators. For details on how to register, see the JaLC website https://japanlinkcenter.org/top/service/service_others.htmlhttps://japanlinkcenter.org/top/service/service_others.html [21].

[21] Japan Link Center, "Other Services," n.d. [Online]. Available: https://japanlinkcenter.org/top/service/service_others.html.

4.5.2 Citing data using DOI

When citing data for which a DOI has been registered in a journal article or other publication, the citation format should follow the submission guidelines of the journal or other publication. However, by specifying the DOI and selecting the citation format in DOI Citation Formatter (https://citation.crosscite.org), citation information for the data corresponding to the citation format can be easily obtained [22]. For more information on how to use this service, see the JaLC website [21].

4.5.3 Description of relevant information

When a dataset for which a DOI has been registered is cited in the content such as journal articles or other data, information about the content and the relationship between the dataset and the content can be described in the "Related information" item of the metadata for the registered DOI for the data [20]. For details on describing metadata, see Appendix 4 and Appendix 5 of the JaLC Content Registration Manual XML Format Guide [Research Data].

4.5.4 How to add or modify data after DOI registration

A dataset for which a DOI has been registered is expected to remain unchanged thereafter, but in the case of dynamically changing data, additions and modifications may be made. In this case, you can choose to either keep the DOI or register a different DOI.

- Here are some examples of possible solutions when continuing using the original DOI after adding data.
  - Example (1):
  Perform version control. Any modifications or changes in data additions, acquisition methods, processing methods, corrections, errors, etc., should be clearly indicated on the landing page as a separate version, allowing access to each version.
  - Example (2):
  The landing page should include a history and details of all modifications and changes in data additions, acquisition methods, processing methods, corrections, errors, etc.
  - Example (3):
  If data are added as necessary during observation, the original DOI is retained, and no version control is applied.

  These examples should be optimized and used according to the characteristics of each field and each data.

- Here are some examples of possible solutions when registering a new DOI.

[22] DOI Registration Agencies, "DOI Citation Formatter," n.d. [Online]. Available: https://citation.crosscite.org/.

[21] Japan Link Center, "Other Services," n.d. [Online]. Available: https://japanlinkcenter.org/top/service/service_others.html.

[20] Japan Link Center, "JaLC Content Registration Manual XML Format Guide [Research Data]," 2023. [Online]. Available: https://japanlinkcenter.org/manual/research_data/index.html.

• Example (1):

Version control is performed, and the version information is clearly indicated in the DOI of the original data and in the metadata of the DOI after the data are added.

• Example (2):

In the case of observation datasets that have been acquired over a long period of time, datasets to which new data have been added should be aliased as a series, divided by a certain unit time and period (e.g., in years), and a DOI should be registered for each unit.

When another DOI is registered, it is recommended to keep the data corresponding to the original DOI, but sometimes it is difficult or not very meaningful to keep the data, such as when a large dataset is corrected as a whole. If the original dataset is no longer accessible for such reasons, the landing page should be retained, and the reason clearly stated.

4.5.5 Merging, dividing, transferring, and discarding data with registered DOI

Data with registered DOIs are expected to remain publicly available for a long period of time, but in some cases, data may be merged, divided, transferred to other organizations, or discarded. In anticipation of such a situation, it is recommended that organizations registering DOIs develop a long-term plan that includes policies for dealing with changes in data release status and for transferring data to other organizations and sharing data with the research community.

When data are merged or divided, a description of the merging or dividing and the DOI after the merging or dividing should be clearly indicated on the landing page of the DOI before the merging or dividing. The relationship with the DOI before the merging or dividing should be also clearly indicated on The DOI landing page after the merging or dividing. (Dividing example: https://doi.org/10.17596/0000314 [23]).

Even if the published data are withdrawn due to data discarding, the DOI landing page should be maintained to continue publishing metadata and ensure access. The fact that the data publication was withdrawn and the reason why should be clearly indicated on the landing page.

[23] JAMSTEC, "NATSUSHIMA NT00-06 Cruise Data," JAMSTEC, 2016. [Online]. Available: https://doi.org/10.17596/0000314.

Chapter 5

Case Studies

National Institute of Polar Research (NIPR), Arctic Data Archive System (ADS)

Features

- Projects implemented within one institution
- Target data: survey and observation data from the Arctic and Antarctic regions
- The manager registers the DOI with JaLC.
- Systematically manage metadata quality

Summary of DOI registration

ADS Metadata Management System (AMS) provides centralized metadata management and data archiving. Metadata and actual data are generated by researchers. When metadata is registered, the system checks it according to the ADS metadata schema, thereby controlling the quality of the metadata.

AMS allows only managers to register DOIs in accordance with the NIPR DOI registration policy. Additionally, a landing page is automatically created based on all registration data and metadata.

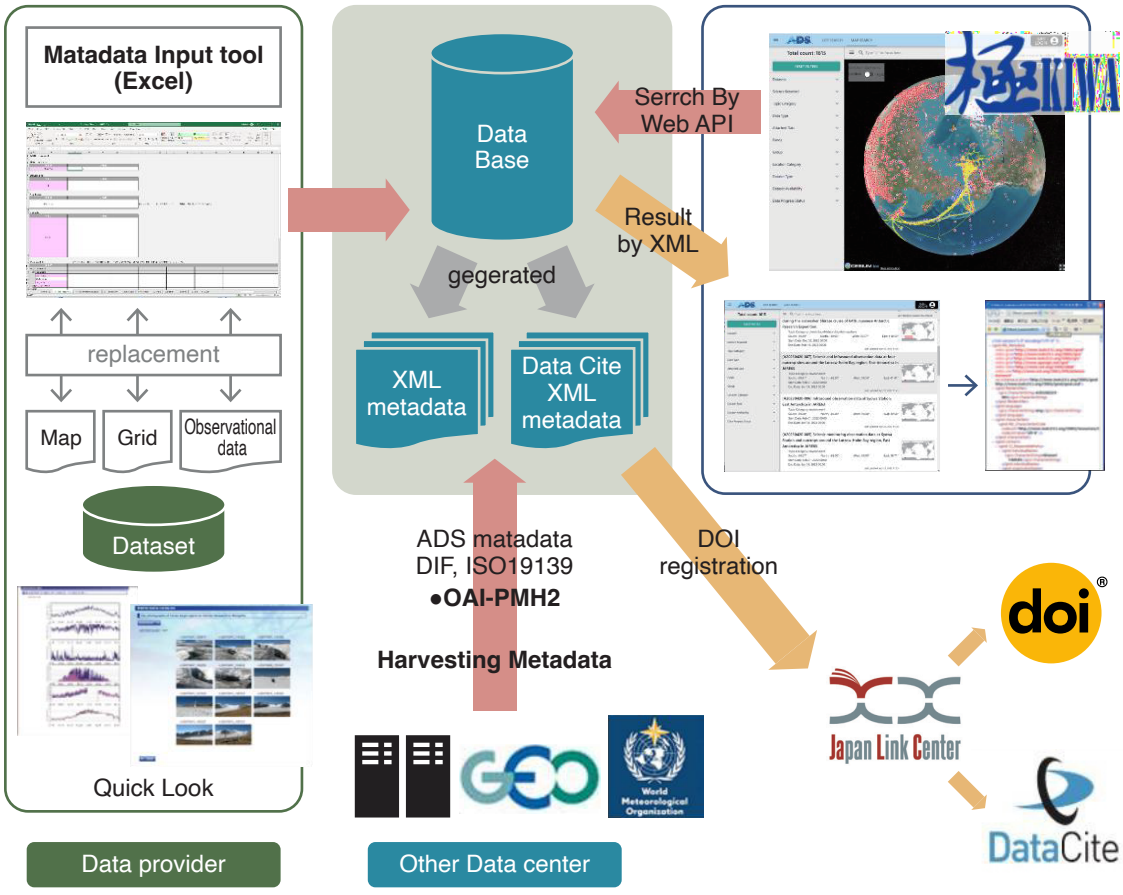


Figure 5-1. Arctic Data Archive System (ADS)

Data Integration and Analysis System (DIAS), a global environmental data integration and analysis platform project

Features

- DIAS is a domain repository for global environmental data, including ground observation data, satellite observation data, weather forecasting models, climate change projection models, and other social data
- DIAS established procedures for accepting, reviewing, and publishing research data not only from the DIAS project but also from other institutions and public data from the government
- Metadata is published under a CC0 license to promote the distribution of data with DOI
- DIAS develops applications that take advantage of DOI, including the Mahalo Button, a system for monitoring dataset usage and sharing it on the dataset landing page

Summary of DOI registration

- Responding to a DOI request from a data creator, DIAS assigns a DOI after agreeing with the data creator about the granularity and other details of the DOI
- In principle, the granularity of a DOI corresponds to a DIAS's dataset ID, but other granularities can be used according to individual requests from a data creator
- Data creators and the DIAS team work together to input the metadata into the DIAS metadata management system, and DIAS manages it as the document metadata of each dataset ID
- The landing page for each dataset is automatically generated from the document metadata
- When a DOI is assigned or updated, the DIAS metadata is automatically converted to the JaLC metadata format and is sent to DataCite via the JaLC API

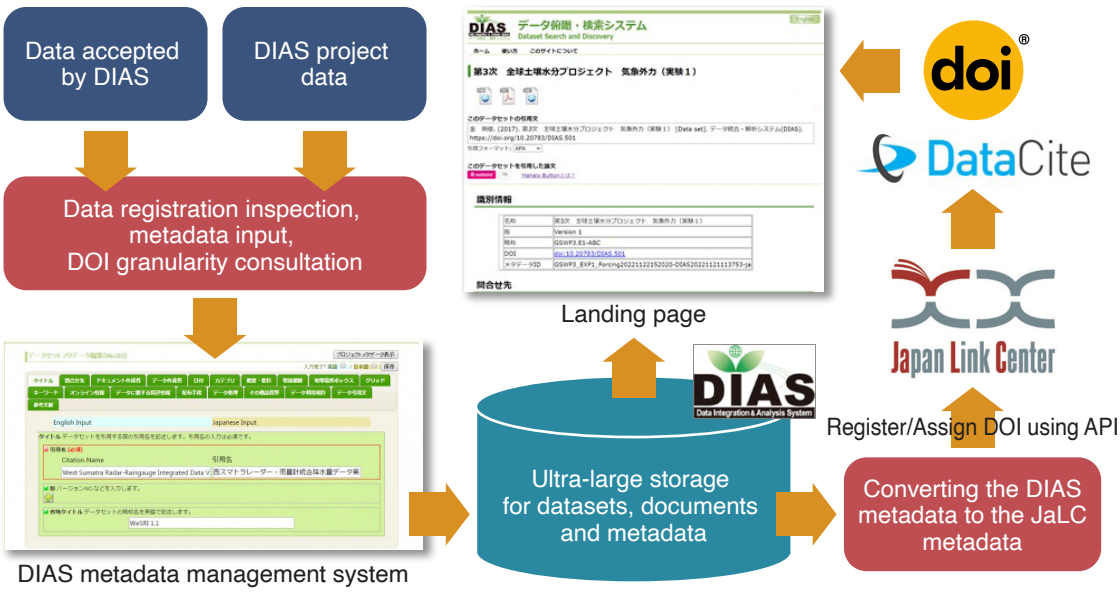


Figure 5-2. DIAS metadata management system

National Institute for Materials Science

Features

- Target data: research data in the field of materials science
- The repository registration system links the data repository for external publications with the in-house research achievement registration system and researcher directory system

Summary of DOI registration

The MDR Registration System was established to register data in the Materials Data Repository (MDR) and to register DataCite DOIs. The “MDR Registration System” automatically retrieves bibliographic information for journal articles to be registered from the institution's internal research achievement registration system and researcher directory system. In addition, data registered in the system can be registered for DataCite DOI from within the system. A link to the MDR is added to the detailed information page of such articles displayed in the institution's researcher directory “SAMURAI” for the accompanying data registered in the MDR through the MDR registration system.

Data registration is performed by researchers (proxy registration is possible), while DOI registration and publication in the repository are performed by MDR operation managers.

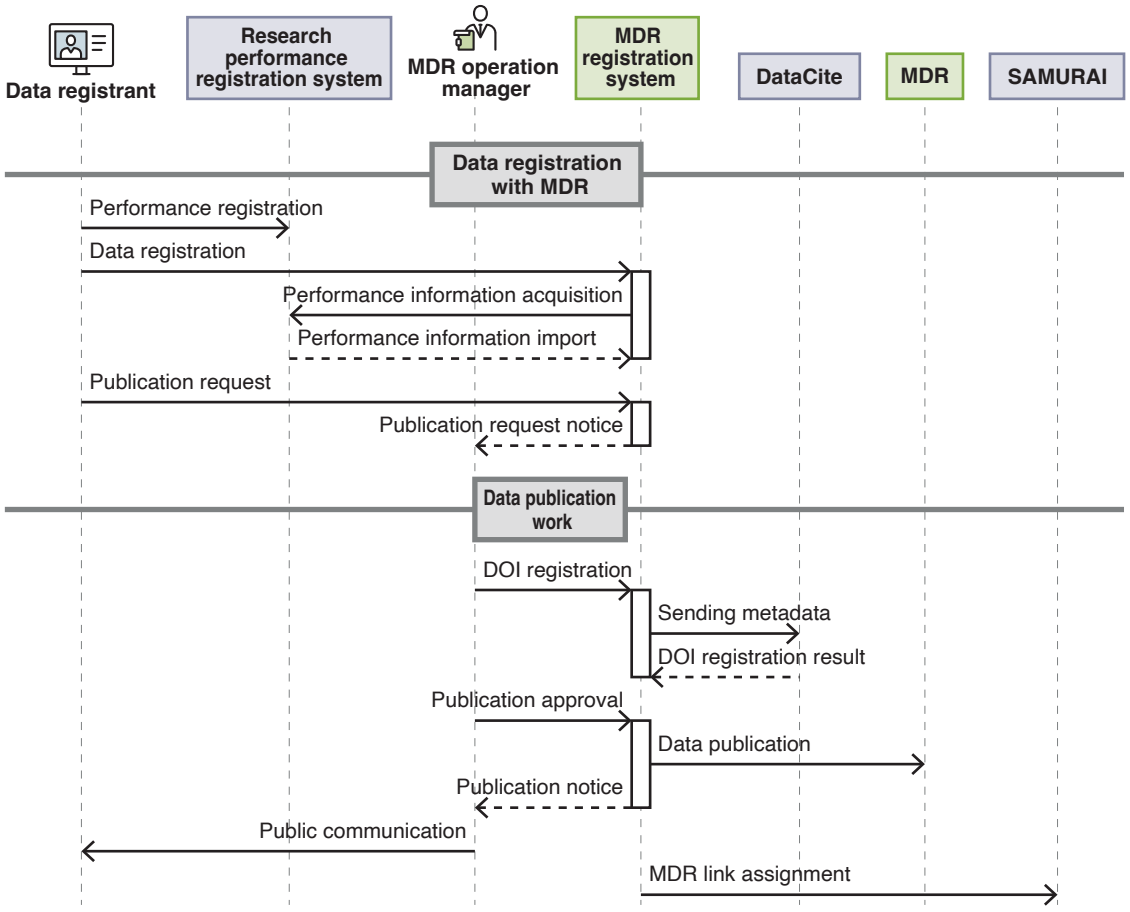


Figure 5-3. Workflow and data flow of DOI registration in Materials Data Repository

RIKEN Center for Brain Science, Neuroinformatics Unit

Features

- Target data: the research data across the entire field of brain and neuroscience. Each dataset is enriched with metadata tailored to specific research areas by using life science ontology and terminology.
- Register and manage DOIs for publicly available data in three research data repositories and databases: the CBS Data Sharing Platform (a research data management platform that collects, shares, and publishes data generated by laboratories within the RIKEN Center for Brain Science), the Brain/MINDS Data Portal (a research data publication system developed by the “Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS)” project), and the INCF Japan Node Platform (a group of databases for each research field constructed through the activities of the INCF Japan Node).
- Register and manage DOIs for publicly available data in three research data repositories and databases: the CBS Data Sharing Platform (a research data management platform that collects, shares, and publishes data generated by laboratories within the RIKEN Center for Brain Science), the Brain/MINDS Data Portal (a research data publication system developed by the “Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS)” project), and the INCF Japan Node Platform (a group of databases for each research field constructed through the activities of the INCF Japan Node).
- All research data published with a DOI undergoes peer review by the professional researcher community to ensure their academic value to the field and the quality of their metadata.

Summary of DOI registration

CBS Data Sharing Platform

Utilize the data publication and peer review workflow functions implemented in the system. After the Repository Committee reviews and approves the data and metadata for the dataset submitted by a researcher for publication, a landing page and metadata for DOI registration are created. DOI registration is conducted indirectly through integration with R2DMS (RIKEN Research Data Management System). Metadata are periodically harvested via the ResourceSync Protocol from NII IRDB (Institutional Repository Database).

Brain/MINDS Data Portal

Research datasets produced by researchers participating in the Brain/MINDS project are collected, registered, and published on the portal site along with metadata under the supervision of the database committee. DOI registration is performed manually by the system administrator using the JaLC web form and JaLC-formatted XML for metadata registration.

INCF Japan Node Platform

The Neuroinformatics Unit continues to maintain and manage the DOIs that have been assigned to the content of the platform that has been operated and constructed by the INCF Japan Node by continuing to operate the publication system.

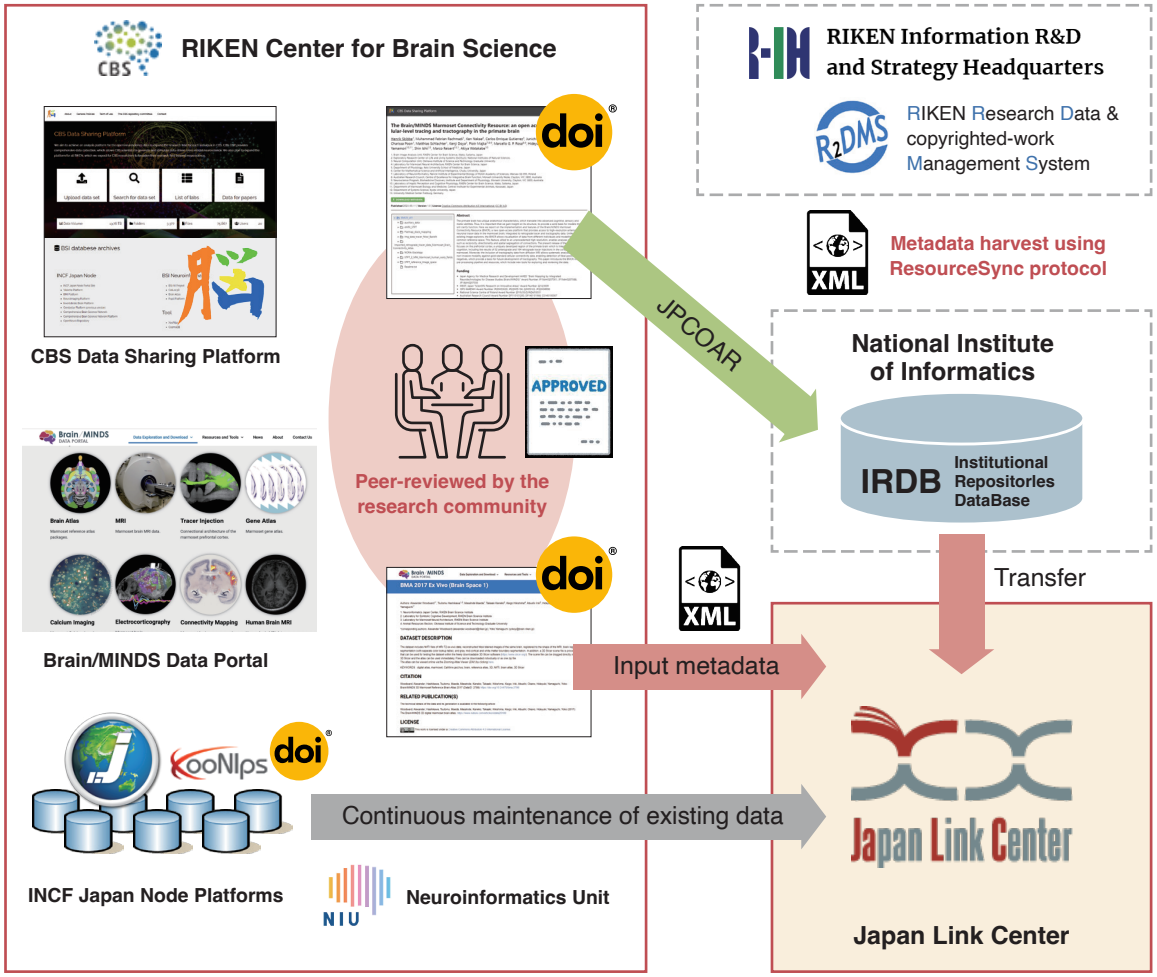


Figure5-4. Flow of DOI registration in the Neuroinformatics Unit

Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Data DOI Information Management System

Features

- Target data: data published by research departments and researchers affiliated with the institution
- Developed internal guidelines for DOI registration
- Established "Data DOI Information Management System" and implemented centralized management
- DOI registration to JaLC is handled by the Data Management Office

Summary of DOI registration

Research departments or researchers create datasets, dataset metadata, and DOI metadata, and make the datasets and dataset metadata publicly available in each database, and also apply for DOI registration to the Data Management Office.

The Data Management Office uses the “Data DOI Information Management System” to manage DOI metadata (DOI numbers, metadata registration and updating), registers it with the JaLC system using the Web API, and publishes landing pages.

To streamline DOI registration process, for some content, the system that manages the content generates XML for registration in the "Data DOI Information Management System." There are also measures in place to prevent broken links when transitioning from the landing page to each database, even if the databases have differe nt life cycles. (DOI example: https://doi.org/10.17596/0000001 [24])

[24] M. Aoyama, "Global Nutrients Dataset 2013," JAMSTEC, 2017. [Online]. Available: https://doi.org/10.17596/0000001.

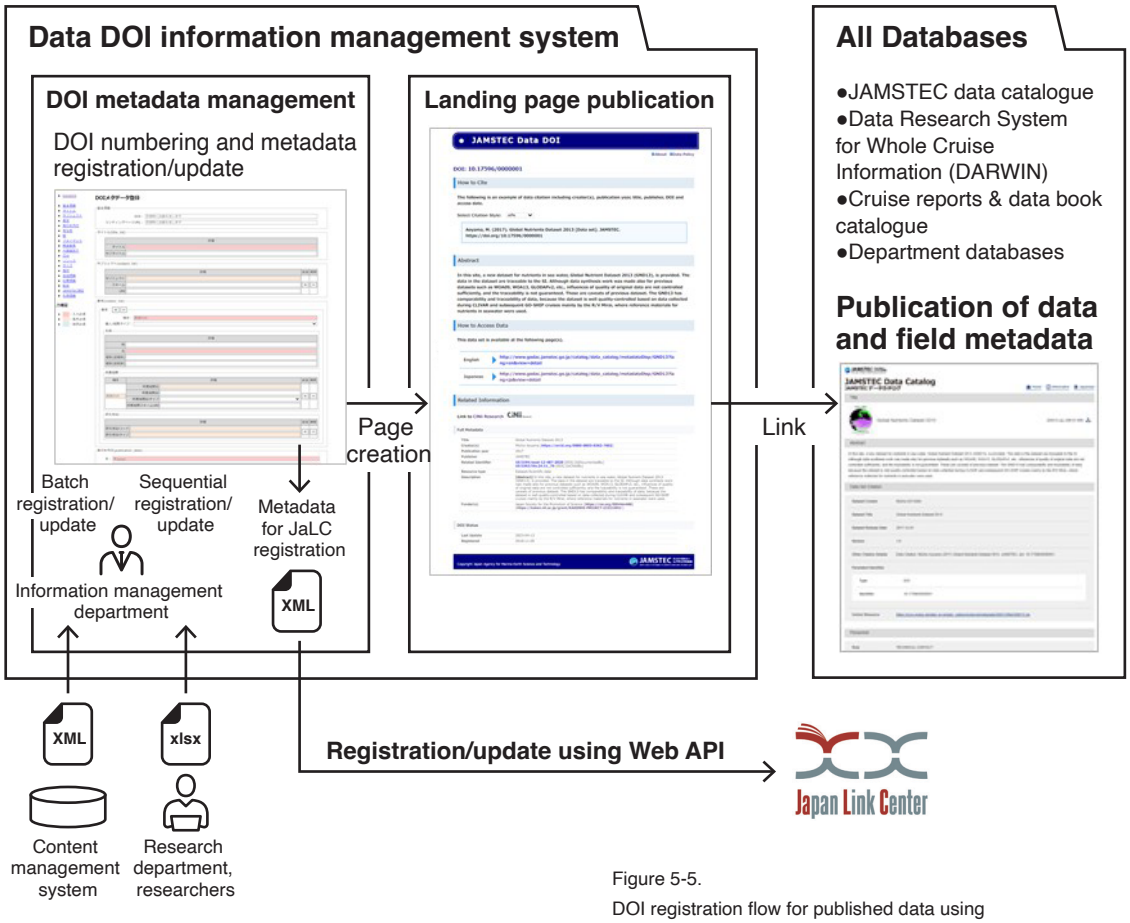


Figure 5-5.
DOI registration flow for published data using
the “Data DOI Information Management System” at JAMSTEC

Japan Agency for Marine-Earth Science and Technology (JAMSTEC), BISMaL: Biological Information System for Marine Life

Features

- Target data: biodiversity datasets published by JAMSTEC in BISMaL
- The datasets published in BISMaL include not only biological sample information obtained during JAMSTEC surveys, but also a large number of data from other institutions in Japan. To distinguish these datasets from JAMSTEC's own published datasets (prefix: 10.17596), they are registered using the DOI prefix "10.48518" assigned to BISMaL.
- The datasets published in BISMaL are linked to the international databases OBIS and GBIF

Summary of DOI registration

BISMaL manager registers DOIs based on the request from data owner.

DOI metadata information is extracted from the metadata of datasets registered in BISMaL, and manually registered using the JaLC system.

The DOI landing page uses the BISMaL dataset description page.

(DOI example: https://doi.org/10.48518/00001 [25])

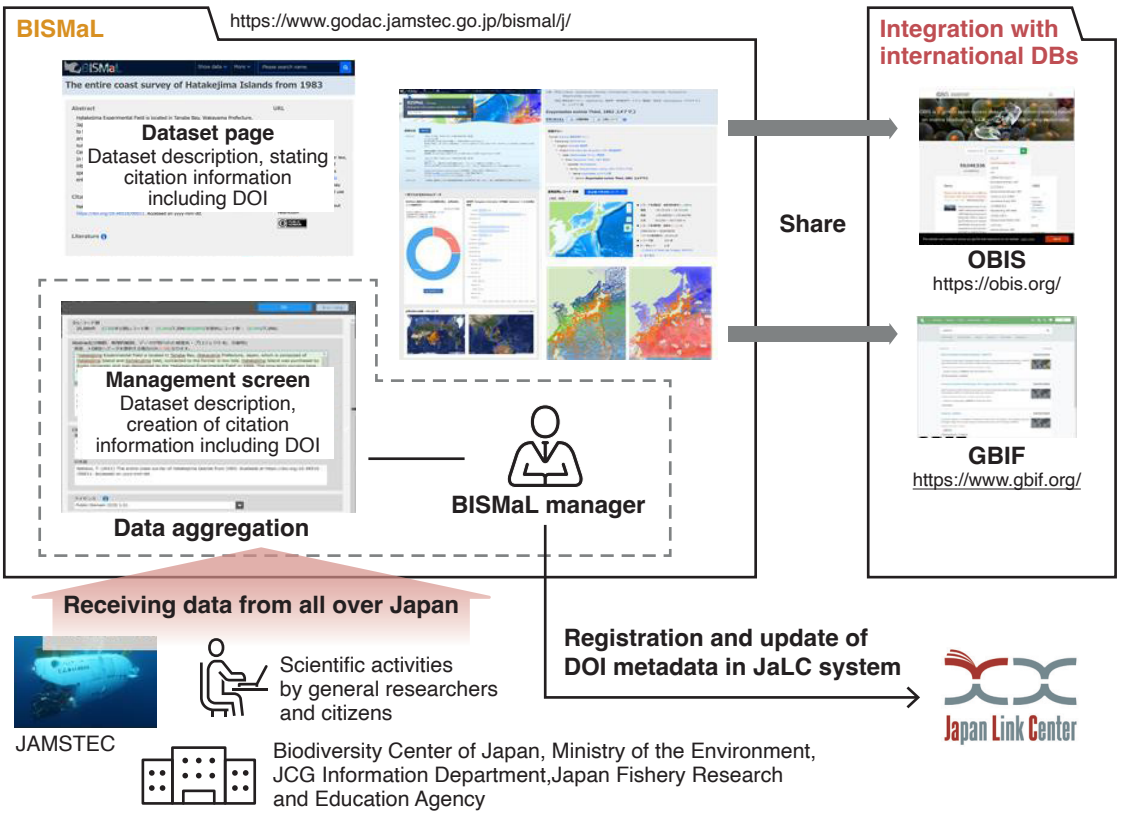


Figure 5-6. Flow of data publishing and DOI registration in BISMaL

[25] JAMSTEC, "JAMSTEC Marine Biological Samples Database," Japan Agency for Marine- Earth Science and Technology, 2016. [Online]. Available: https://doi.org/10.48518/00001.

Historiographical Institute, the University of Tokyo

Features

- As part of the JSPS Program for Constructing Data Infrastructure for the Humanities and Social Sciences (FY 2018-2022), the JSPS is conducting DOI registration for research data in the humanities.
- In addition to metadata about historical records held by the Historiographical Institute, the University of Tokyo, the institute also registers DOIs for metadata published by collaborative institutions, playing a central role in the utilization of research data and the distribution of academic information in the humanities.

Summary of DOI registration

A “DOI Management System” is established to aggregate and convert metadata for DOI registration. In addition to the API integration function that acquires metadata about historical records provided by the Historiographical Institute's Historical Information Processing System (SHIPS), the system has a function to harvest metadata provided by collaborative institutions such as the Kanagawa Prefectural Kanazawa-Bunko Museum. By using the data aggregated and converted by the DOI management system to register DOIs with JaLC and sending metadata including the obtained DOIs to JDCat, a comprehensive data catalog for the humanities and social sciences, the institution supports the utilization of research data and the distribution of academic information in the humanities.

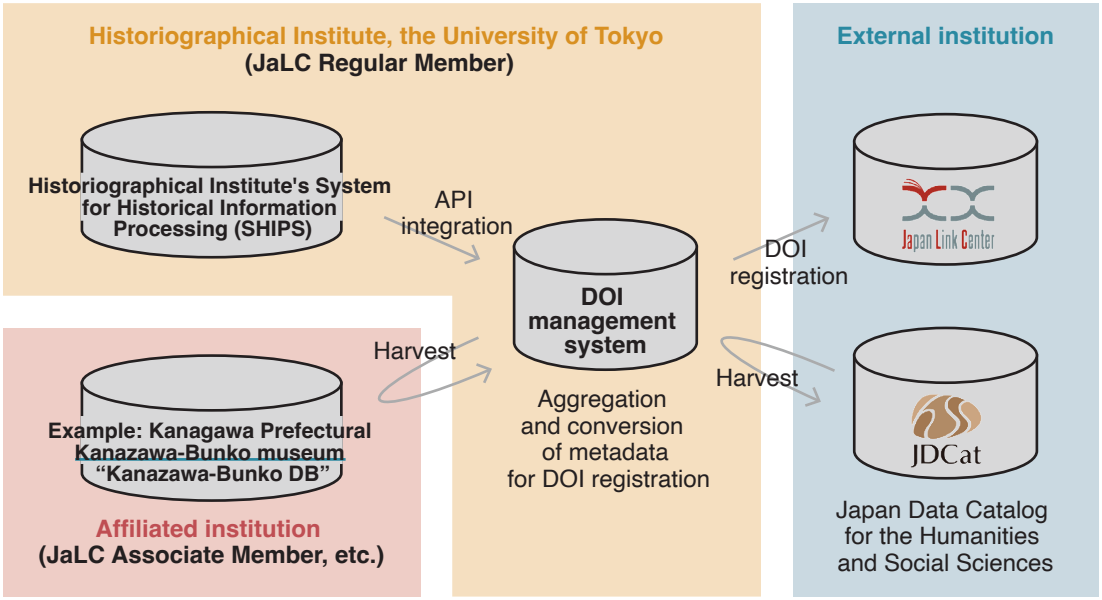


Figure 5-7. Schematic diagram for the Historiographical Institute of the University of Tokyo

National Institute for Environmental Studies (NIES): Global Environmental Database (GED)

Features

- GED collects, maintains, and provides data and results from research in the natural and social sciences field on global warming and other global environmental issues.
- NIES became a JaLC member and began registering DOIs for the research data since 2016.
- A set of metadata items are optimized for each major data category (fixed point observations, mobile observations, and models). Further customization is allowed according to the needs for each dataset.

Summary of DOI registration

- The Center for Global Environmental Research (CGER) in the Earth System Division (one of the research units of NIES) is in charge of supporting metadata creation by data providers, creating metadata page, XML file for DOI registration, and the data publication page, while the Environmental Information Division (one of the administration departments of NIES) is in charge of DOI registration through JaLC and the creation and operation of the DOI landing page (Figure 5-8). In consideration of the persistence of DOIs, DOI registration and the landing page management are assigned to an administration department, under the condition that research units are restructured for each mid- to long-term plan of the institute.
- The web application GERDaMS (Global Environmental Research Data Management System), which supports efficient management of research data, data sharing among research teams, metadata creation, licensing, versioning, and DOI registration for smooth data publication is being developed and operated along with GED (Figure 5-9).

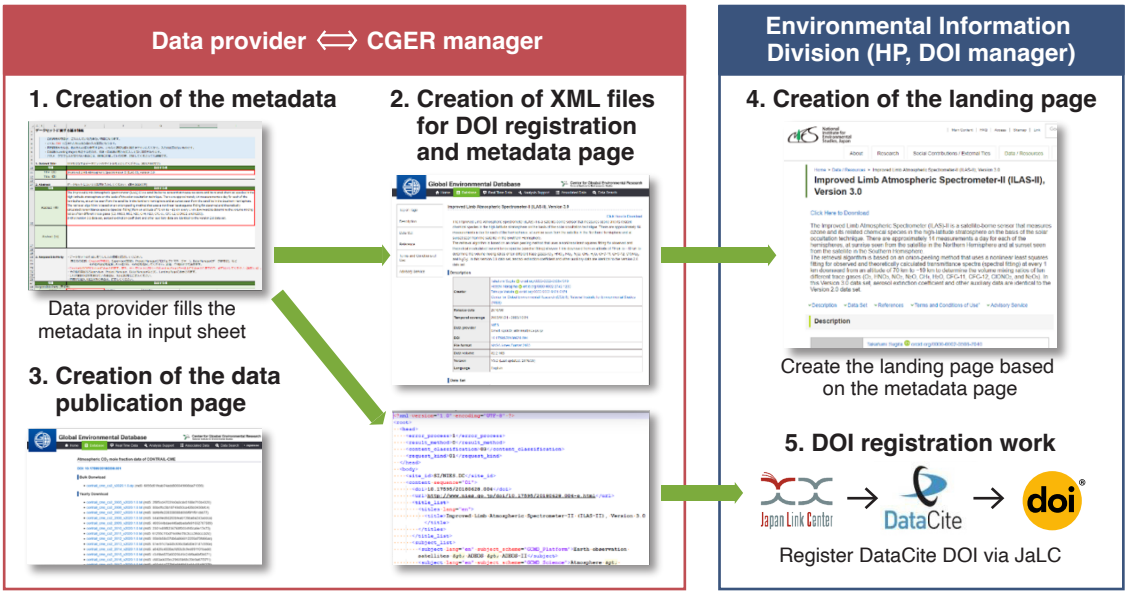


Figure 5-8. Workflow for registering DOIs for research data in GED

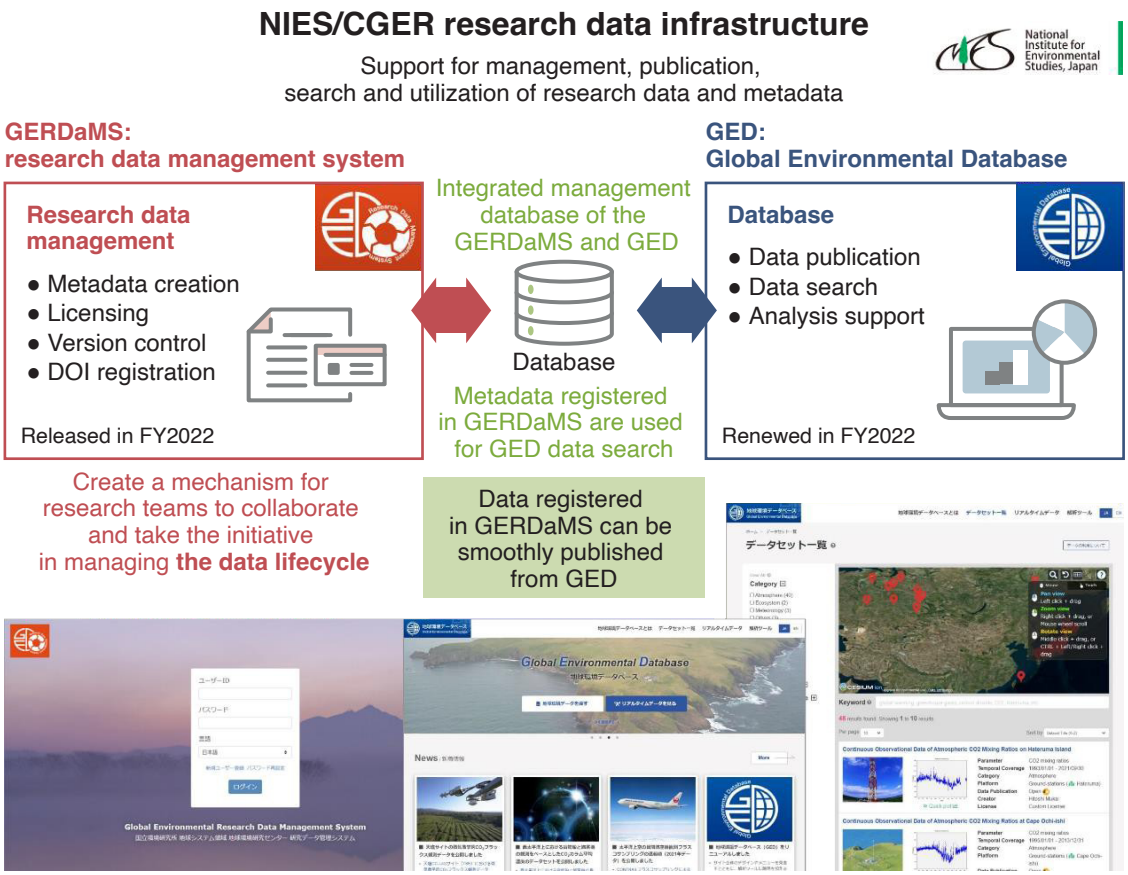


Figure 5-9. NIES/CGER research data infrastructure

Institute for Space–Earth Environmental Research, Nagoya University, Tokai National Higher Education and Research System

Features

- Target data: Various research data (ground-based and satellite observation data, catalog lists, and model/simulation data) related to the research community.
- Using a tool, researchers prepare an XML file including those needed for DOI registration on the JaLC system and some additional metadata. From that XML file, an XML file (in the JaLC format) for DOI registration and an HTML file for the landing page are generated using another tool developed in Java script languages.

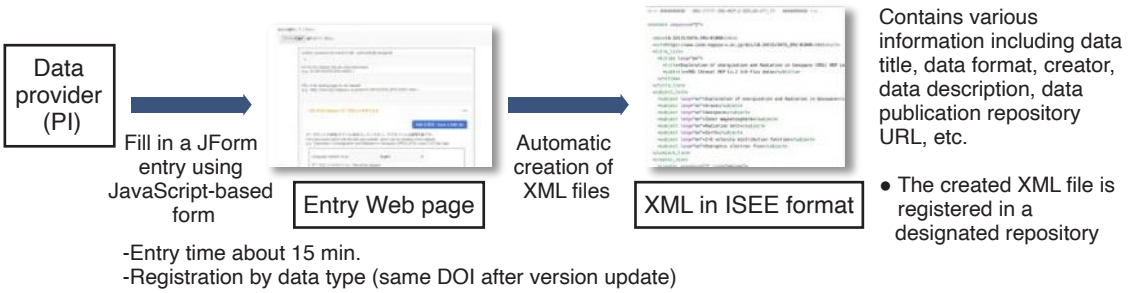
Summary of DOI registration

- Researchers generate an XML file for DOI application using a Javascript-based webform tool, which was developed by the institute. Another software is then used to convert the XML file to an XML file for DOI registration on the JaLC system and an HTML file for its landing page.
- After being reviewed by staff at the Institute, the DOI application in XML is submitted to JaLC. The landing page is published as soon as the DOI registration is completed.

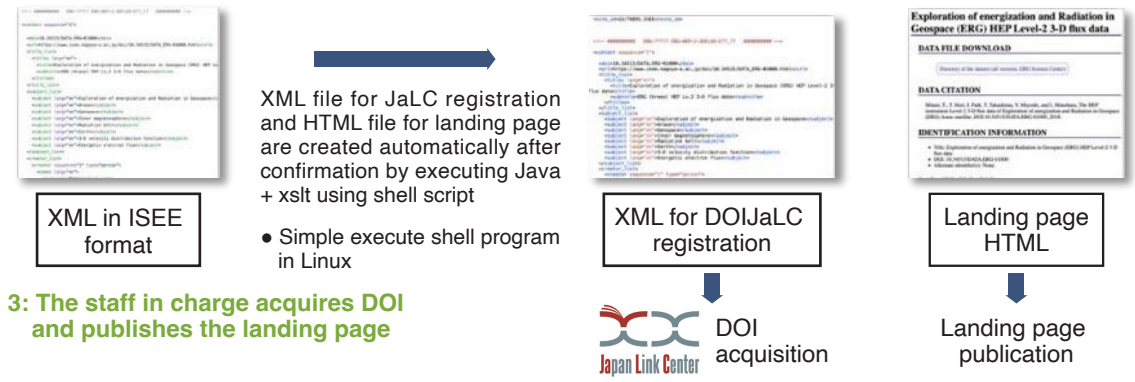
Institute for Space–Earth Environmental Research (ISEE), Nagoya University, Tokai National Higher Education and Research System

DOI assignment flow

1: XML file for registration is created by the data provider and registered in the repository



2: The staff in charge confirms and converts to XML file for JaLC registration and HTML file for landing page



3: The staff in charge acquires DOI and publishes the landing page

Figure 5-10. Workflow of DOI registration by ISEE

References

[1] Japan Link Center Experimental Project for DOI Registration for Research Data, "Guidelines for Registering DOIs for Research Data,"
10/20/2015. [Online] Available: https://doi.org/10.11502/rd_guideline_ja.

[2] The DOI Foundation, "DOI® HANDBOOK: DOI," 4 2023. [Online]. Available:
https://doi.org/10.1000/182.

[3] Japan Link Center Steering Committee, "What is the Japan Link Center? Its origins and basic policy," 7/28/2014. [Online] Available: https://doi.org/10.11502/JaLC_policy.

[4] Crossref, "You are Crossref: Crossref," n.d. [Online]. Available: https://www.crossref.org/.

[5] DataCite, "DataCite-Connecting Research, Advancing Knowledge," n.d. [Online]. Available: https://datacite.org/.

[6] Japan Link Center, "Overview of JaLC: Japan Link Center (JaLC)," n.d. [Online]. Available: https://japanlinkcenter.org/top/about/index.html.

[7] Japan Link Center, "Membership type and annual membership fee: Japan Link Center (JaLC)," n.d. [Online]. Available: https://japanlinkcenter.org/top/admission/member_type.html.

[8] Japan Link Center, "DOI Registration: Japan Link Center (JaLC)," n.d. [Online]. Available: https://japanlinkcenter.org/top/service/service_doi.html.

[9] Japan Link Center, "Japan Link Center (JaLC) Release Summary," 1/26/2022. [Online]. Available: https://japanlinkcenter.org/top/doc/ReleaseNotes_20220126_suppl.pdf.

[10] DataCite, "DataCite Metadata Schema 4.4," 30 5 2021. [Online]. Available:
https://schema.datacite.org/meta/kernel-4.4/.

[11] Council for the Promotion of Integrated Innovation Strategy, Cabinet Office, Government of Japan, "Basic Approach to Management and Utilization of Publicly Funded Research Data," 2021. [Online]. Available: https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/sanko1.pdf.

[12] Junichi Onami, Shigeru Yatsuzuka, Tomoe Nobusada, Mari Minowa, Nobutaka Mitsuhashi, Hideki Hatanaka, "The FAIR Principles as a Standard for Data Sharing," Bioscience Database Center, 2018. [Online]. Available: https://doi.org/10.18908/a.2018041901.

[13] Japan Link Center, "JaLC Content Registration Manual Web Interface Version [Research Data]," 2023. [Online]. Available: https://doi.org/10.11502/web_interface_manual_research_data.

[14] Japan Link Center Steering Committee, "Japan Link Center Terms of Participation," 2023. [Online] Available: https://japanlinkcenter.org/top/doc/jalc_sankakiyaku.pdf.https://japanlinkcenter.org/top/service/service_doi.html.

[15] Cabinet Office, Japan Open Science Promotion Committee, "Guidelines for Development of Data Policies for National Research Institutes," 6/29/2018. [Online]. Available: https://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.html.

[16] Masahito Nose, Kazuki Ohmukai, Mizuho Owashi, Nanako Takahashi, & Yasuhiro Murayama, "How to assign DOIs to research data? A five minute guide to assigning DOIs to research data," Research Data Citation Subcommittee of the Research Data Utilization Forum (RDUF), 2020. [Online]. Available: https://doi.org/10.11502/rduf_rdc_doileaflet.

[17] Yasuhiro Murayama, Masahito Nose, Yukinobu Koyama, Hideaki Takeda, Koji Zettsu, Takenari Kinoshita, Toshihiko Iyemori, and Takashi Watanabe, "Proposal of a new DOI citation method for dataset generated from other datasets," 138th SGEPSS General Meeting and Lecture, Society of Geomagnetism and Earth, Planetary and Space Sciences, 2015.

[18] S. Murakami, S., M. Yamada, A. Yamazaki, K. McGouldrick, Y. Yamamoto, G.L. Hashimoto, "Venus Climate Orbiter Akatsuki UVI Calibrated Data PDS3 dataset," 11 11 2017. [Online]. Available: https://doi.org/10.17597/ISAS.DARTS/VCO-00003.

[19] Zenodo, "Frequently Asked Questions: Versioning," n.d. [Online]. Available: https://help.zenodo.org/faq/.

[20] Japan Link Center, "JaLC Content Registration Manual XML Format Guide [Research Data]," 2023. [Online]. Available: https://japanlinkcenter.org/manual/research_data/index.html.

[21] Japan Link Center, "Other Services," n.d. [Online]. Available: https://japanlinkcenter.org/top/service/service_others.html.

[22] DOI Registration Agencies, "DOI Citation Formatter," n.d. [Online]. Available: https://citation.crosscite.org/.

[23] JAMSTEC, "NATSUSHIMA NT00-06 Cruise Data," JAMSTEC, 2016. [Online].Available: https://doi.org/10.17596/0000314.

[24] M. Aoyama, "Global Nutrients Dataset 2013," JAMSTEC, 2017. [Online]. Available: https://doi.org/10.17596/0000001.

[25] JAMSTEC, "JAMSTEC Marine Biological Samples Database," Japan Agency for Marine-Earth Science and Technology, 2016. [Online]. Available: https://doi.org/10.48518/00001.

[26] METADATA 20/20, "LINKS TO BEST PRACTICES AND GUIDELINES OF THE SCHOLARLY COMMUNICATIONS COMMUNITY," n.d. [Online]. Available: https://metadata2020.org/resources/metadata-best-practices/.

[27] METADA 20/20, "THE METADATA 2020 PRINCIPLES," n.d. [Online]. Available: https://metadata2020.org/resources/metadata-principles/.

[28] Crossref, "Documentation: Datasets," n.d. [Online]. Available: https://www.crossref.org/documentation/principles-practices/datasets/.

[29] DataCite, "Best Practices for DOI Registration," 2023. [Online]. Available: https://support.datacite.org/docs/best-practices-for-datacite-members.

[30] ORCiD, "ORCiD Connecting research and researchers," n.d. [Online]. Available: https://orcid.org/.

[31] National Institute for Materials Science (NIMS), "MDR Schema," n.d. [Online]. Available: https://github.com/nims-dpfc/mdr-schema?tab=readme-ov-file#mdr-schema.

[32] Japan Aerospace Exploration Agency, "ISAS Data Policy," 18 4 2018. [Online]. Available: https://www.isas.jaxa.jp/researchers/data-policy/.

[33] Y. Murayama, "Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaska," 2012. [Online]. Available: https://doi.org/10.17591/55838dbd6c0ad.

[34] Japan Link Center, "Services: DOI Registration," [Online]. Available: https://japanlinkcenter.org/top/service/service_doi.html. [Accessed 14 5 2024].

Appendices

Appendix 1 Terms

Creative Commons (CC)

An international project and non-profit organization that operates it. The project defines and publishes various license formats which allow authors to publish and permit the use of their works for the purpose of reuse and promotes the utilization and dissemination of these formats. The multiple licenses that have been developed are called Creative Commons licenses and are widely used by the public, primarily in the academic field. (https://creativecommons.jp/)

Crossref

Largest DOI registration agency (RA). Primarily deals with journal articles. A global community infrastructure that makes it easy to find, evaluate, and reuse all types of research subjects through a number of essential services for research communication. (https://www.crossref.org/)

DataCite

A DOI registration agency (RA) that primarily deals with research data. It is a global non-profit organization that provides DOIs for mainly research data and all other research outputs. It was established in 2009 by European and American universities and libraries, with the German National Library of Science and Technology (TIB) serving as its secretariat.
(https://www.datacite.org/)

deposit

The act of registering electronic data, with the purpose of depositing it safely.
Used in the sense as the metadata and text converted to XML are deposited into the JaLC system.

DOI

Digital Object Identifier. Refer to the DOI Handbook . A permanent identifier assigned to a document on the Internet.
A typical DOI name is a string like 10.1246/nikkashi1898.1.1. Here, 10.1246 (the prefix) is a directory identifier assigned by the DOI Foundation. The slash followed by nikkashi1898.1.1 (the suffix) is an arbitrary ID assigned by the DOI registration agency (in this example, the Chemical Society of Japan). When actually searching by DOI in a browser, simply add the DOI string after the URL "http://doi.org" as follows:http://doi.org/ https://doi.org/10.1241/johokanri.56.881

DOI Foundation

Manages the Digital Object Identifier (DOI) system on behalf of the DOI registration agencies (RAs) who serve their respective communities. (https://www.doi.org/)

DOI prefix

Directory identifier assigned by the DOI Foundation. See DOI.

DOI registration

Register a DOI for the content and registering (depositing) it with the RA along with the URL and metadata.

DOI resolution, DOI resolver

DOI resolution means entering a DOI and obtaining the URL of a Web page that represents the content itself or its metadata. The network service that resolves DOIs is called a DOI resolver.

DOI suffix

An identifier for the directory registered by the DOI registrant. See DOI.

FAIR Principles

A guideline for open data and open access that was developed in 2011 to increase the publication and reusability of research data. Consists of the four initial letters: Findable, Accessible, Interoperable, and Reusable. The first paragraph states that "F1. (meta) data shall have a globally unique and persistent identifier (ID)." (https://doi.org/10.18908/a.2019112601)

granularity

In this guideline, the term means the size of the unit (piece) of research data for which DOI is registered. Increasing the granularity of data •datasets will result in fewer DOIs being registered.

issuing and returning prefixes

The prefixes are issued to each RA by the DOI Foundation and assigned to the institution or project unit that wishes to register the DOI.

metadata

Metadata is not the data itself, but attribute values and additional information of the data that can be searched by information search systems. In books, journal articles, etc., it is sometimes called bibliographic information. For example, for documents, the author's name, title, date of publication, and related keywords are commonly used. In addition to this, for research data, data format, usage license, etc. are also applicable.

metadata schema

Refers to the combination of metadata data structure and the rules for describing it. In the field of information retrieval, it is standard practice to express metadata schema in XML language (e.g. ISO 19115, etc.), which has the advantage of being highly readable.

multiple resolution

Typically, a DOI is linked with only one Web page, and the DOI resolves to that page. However, when you want to link a DOI to multiple web pages, such as when the same content has multiple publication destinations, it can be linked to an intermediate page that displays the URLs of all web pages linked to the DOI as multiple resolution.

Persistent Identifier (PID)

A DOI is a global, persistent identifier that is publicly available and searchable on the Internet. See also the FAIR Principles.

prefix

Refers to the DOI prefix. See DOI prefix.

RA

Abbreviation for Registration Agency. Means DOI registration agency.

registrant

An organization that registers DOIs for electronic data such as journal articles, articles, and journals.

repository

An information system that accumulates and provides journal articles, research data, and other content. There are institutional repositories that universities and research institution accumulate and provide content produced by affiliated researchers, subject based repositories that accumulate and provide research results in specific fields, and data repositories that specialize in accumulating data rather than documents.

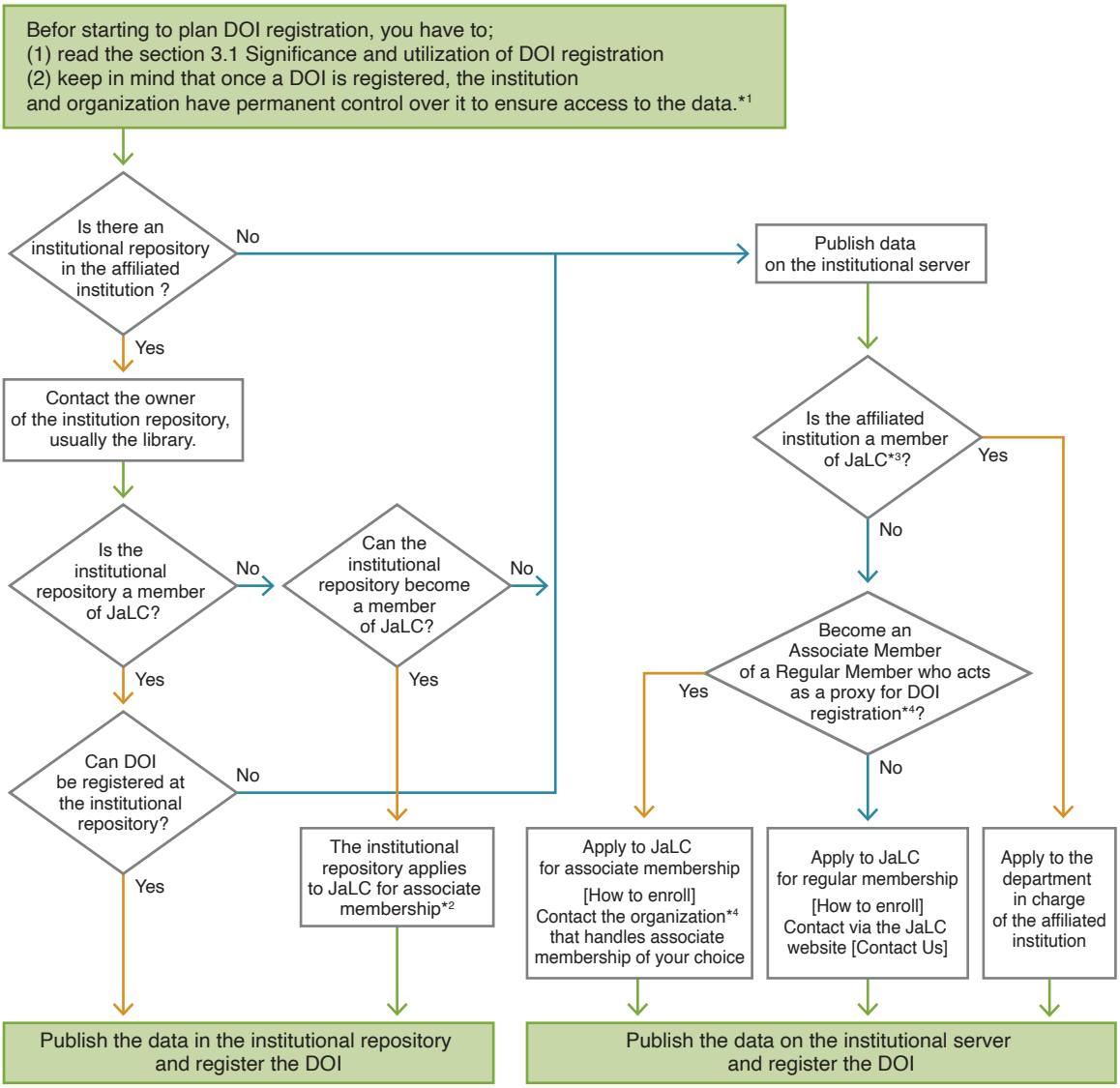
resources

An object that can be named and described, such as a book, digital content, or a person. [Source: National Diet Library website (https://www.ndl.go.jp/jp/dlib/standards/meta/glossary.html)]

suffix

Refers to the DOI suffix. See DOI suffix.

Appendix 2 JaLC enrollment flowchart



*1 If multiple institutions wish to register DOIs for a project, etc., DOI registration will be considered on an institution-by-institution basis. For details, refer to Chapter 3.3, "Points to note when registering DOI."

*2 Assumed to be a JaLC Associate Member of NII. If not applicable, the member becomes another associate or regular member. https://support.irdb.nii.ac.jp/ja/application/j alc

*3 See here for a list of Japan Link Center members https://japanlinkcenter.org/top/doc/listofm embers.pdf

*4 4 See here for a list of organizations handling associate membership https://japanlinkcenter.org/top/admission/ member_type.html

In addition to registering DOIs with JaLC, JaLC offers the registering DOIs for research data with DataCite as an optional service. An overview of each service is shown in Table 7-1.

	JaLC	DataCite (JaLC)	DataCite (Direct Member) *reference
Required membership type	JaLC member (Regular Member, Associate Member)	JaLC member (Regular Member, Associate Member) + DataCite JaLC consortium	DataCite Direct Member
Cost	JaLC annual membership fee	JaLC annual membership fee + DataCite institutional fee 500 Euro (*1) + DOI registration fee (pay-per-use)	DataCite annual membership fee 2000 Euro + DataCite institutional fee 500 Euro (*1) + DOI registration fee (pay-per-use)
Where to register and how to utilize DOI	[Registration] JaLC [DOI utilization] • JaLC content search • Services that JaLC works with (CiNii, ORCID, etc.)	[Registration] JaLC+DataCite [DOI utilization] • JaLC content search • Services that JaLC works with + • DataCite Commons • Services that DataCite works with (re3data, Data Citation Index, etc.)	[Registration] DataCite [DOI utilization] • DataCite Commons • Services that DataCite works with (re3data, Data Citation Index, etc.)
Metadata schema for registration	JaLC scheme	JaLC scheme (*2) ➡ Convert to DataCite schema and register via JaLC. *2 Some elements differ between JaLC and DataCite. Reference: Research Data Deposit XML List of metadata items	DataCite scheme
JaLC support	Yes	Yes	No
	⬇	⬇	⬇
Which service should I choose? For example….	• Content intended primarily for domestic data usage • Metadata mainly in Japanese • Number of DOIs to be registered: tens of thousands per year	• Content expected wide usage both domestic and oversea • Metadata in English • Number of DOIs to be registered: tens of thousands per year	• Content intended primarily for the use of overseas (without emphasis on domestic usage) • Content that wish to contain description from the DataCite schema elements • Number of DOIs to be registered: over 100,000 per year

*1 For non-profit organizations. For profit/commercial organizations, see https://datacite.org/fee-model/#Multiplier-for-For - profit-Organizations

* For information on DataCite DOI registration as an optional JaLC service, please refer to the Japan Link Center Terms of Participation (Article 17 DataCite Service). https://japanlinkcenter.org/top/doc/jalc_sankakiyaku.pdf (in Japanese only)
For more information on fees, refer to the "Japan Link Center Guide to Annual Membership Fees and Fees for External Services" https://japanlinkcenter.org/top/doc/jalc_FeeModel.pdf (in Japanese only)

* For more information, please contact the JaLC Support Center. For contact: info@japanlinkcenter.org

* (Reference) It is also possible to become a Direct Member of DataCite directly without JaLC service. For more information, see https://datacite.org/become-a-member/.

* Note: Based on information from the DataCite website as of May 22, 2024. Please be sure to refer to the latest information regarding membership fee models, etc.

Research Data Utilization Forum (RDUF)
Subcommittee for Promoting DOI Registration for Research Data

Chair Tomoko Shirai (National Institute for Environmental Studies)

Committee Members
Ken Ebisawa (ISAS, JAXA)
Takuya Kadohira (NIMS)
Asanobu Kitamoto (National Institute of Informatics)
Takako Takai (Japan Agency for Medical Research and Development)
Hideaki Takeda (National Institute of Informatics)
Ryuichi Takebe (Japan Society for Information Science and Technology)
Ritsuko Nakajima (Japan Science and Technology Agency)
Hideya Nakanishi (National Institute for Fusion Science)
Kazuhiro Hayashi (National Institute of Science and Technology Policy, MEXT)
Shosuke Hayashi (Graduate School of Science, Kobe University / Meteorological Society of Japan)
Kazuyo Fukuda (Japan Agency for Marine-Earth Science and Technology)
Yoko Fukuda (National Institute for Environmental Studies)
Yasuyuki Minamiyama (National Institute of Informatics)
Nodoka Mimura (Japan Science and Technology Agency)
Yasuhiro Murayama (NICT Knowledge Hub, National Institute of Information and Communications Technology)
Shigeru Yatsuka (Biotechnology Center, National Institute of Technology and Evaluation)

Secretariat of Japan Link Center
Toshiaki Uematsu (Japan Science and Technology Agency)
Runa Kobayashi (Japan Science and Technology Agency)
Kohei Sakai (Japan Science and Technology Agency)
Satoko Fujisawa (Japan Science and Technology Agency)

Period of activity
November 2021 - June 2024

[Institutions participating in Case Studies]
- Research Organization of Information and Systems, National Institute of Polar Research
- Data Integration & Analysis System Program (DIAS)
- National Institute for Materials Science
- RIKEN Center for Brain Science
- Japan Agency for Marine-Earth Science and Technology (JAMSTEC)
- Historiographical Institute, the University of Tokyo
- National Institute for Environmental Studies
- Institute for Space–Earth Environmental Research, Nagoya University, Tokai National Higher Education and Research System

Guidelines for Registering DOIs for Research Data

June 3, 2024

Subcommittee for Promoting DOI Registration for Research Data,
Research Data Utilization Forum (RDUF)
Japan Link Center

5-3, Yonbancho, Chiyoda-ku, Tokyo, 102-8666 Japan
e-mail: jalc-contact@jst.go.jp
https://japanlinkcenter.org/top/english.html

Cite as:
“Guidelines for Registering DOIs for Research”,
Subcommittee for Promoting DOI Registration for Research Data,
RDUF, Japan Link Center, retrieved [date],
https://doi.org/10.11502/rd_guideline_2e_en