

Guidelines for Registering DOIs for Research Data

October 20, 2015

Japan Link Center Joint Steering Committee

Table of Contents

1. Introduction.....	- 2 -
2. DOI registration and JaLC.....	- 2 -
2.1. DOI registration and an overview of JaLC	- 2 -
2.2 JaLC metadata	- 4 -
3. Guidelines for registering a DOI for research data	- 5 -
3.1 Workflow	- 5 -
3.1.1 Life cycle of data, data creator, administrator, and institution	- 5 -
3.1.2 Membership of JaLC	- 7 -
3.1.3 Principles of prefix assignment	- 9 -
3.2 Research data to be registered	- 9 -
3.2.1 Principles of research data to be registered	- 9 -
3.2.2 Quality of research data	- 10 -
3.2.3 DOI coordination between data repositories	- 10 -
3.3 Guarantee of continued access.....	- 10 -
3.3.1 Data held by time-limited projects.....	- 10 -
3.3.2 Removing data after registering a DOI	- 11 -
3.3.3 Data management policy in research activities	- 11 -
3.4 Granularity of DOI registration.....	- 11 -
3.4.1 Basic concept.....	- 11 -
3.4.2 Factors in deciding granularity	- 11 -
3.4.3 Examples of granularity	- 12 -
3.4.4 Addition and modification of data after DOI registration	- 13 -
3.4.5 Relationship with the quantity of data and the number of files.....	- 14 -
3.4.6 Assigning a suffix	- 14 -
3.5 DOI landing page.....	- 15 -
3.6 Establishing your own DOI registration policy.....	- 16 -
4. Case studies.....	- 17 -
5. Bibliography.....	- 22 -
Glossary of Terms.....	- 23 -

1. Introduction

Japan Link Center (JaLC) was founded as a DOI registration agency (RA) to collect, circulate and promote the use of academic materials produced in Japan. Since being authorized as an RA in March 2012, JaLC has been mainly registering DOIs for journal articles (academic papers). On releasing a new system in December 2014, JaLC expanded its content categories for DOI registration to journals (serial publications), research reports, books, research data, and academic materials (e-learning resources), along with journal articles.

Various issues related to the process of registering DOIs for research data are still being discussed across the world and Japan lacks experience of data registration. Before JaLC started registering DOIs for research data, it undertook an experimental project with participants to establish new operation flows for registering DOIs for research data in Japan. Through the project, we have clarified and addressed issues related to DOI registration, worked to establish a stable operation method, and examined how to utilize the registered DOIs.

Each participant has also tested the DOI registration processes. These guidelines are developed based on knowledge obtained through this project and discussions in meetings. In developing these guidelines, we hope that they will be used as a baseline and reference when research institutions start registering DOIs for their research data.

2. DOI registration and JaLC

2.1. DOI registration and an overview of JaLC

The DOI system stores IDs (DOIs) that are assigned to individual content and their corresponding URLs (locations) as pairs and returns a URL in response to a DOI query. When the URL of a digital object is changed, the pair information is updated to guarantee continued access. A DOI starts with “10.” and is separated by the “/” character; e.g., “10.1241/xxx-oo-oo.” The prefix preceding the “/” character is a unique code assigned to the owner of the object by the RA. The suffix following the “/” character can be chosen by registrants. When the address <http://doi.org/DOI> is accessed after DOI registration, the original URL is resolved and as a result users can access the object.

To maintain the system’s consistency, the relevant organizations form a three-layer structure: the International DOI Foundation (IDF), which governs the

(*)

Notes in the right margin of pages show the distinction between: [Mandatory], [Recommended], and [Example] in this document.

[Mandatory] shows that the procedure must be observed, [Recommended] shows that the procedure can be observed, and [Example] shows that the description is

DOI system, registration agencies (RAs), and DOI registrants. JaLC is an RA and its members are DOI registrants. (Figure 2-1)

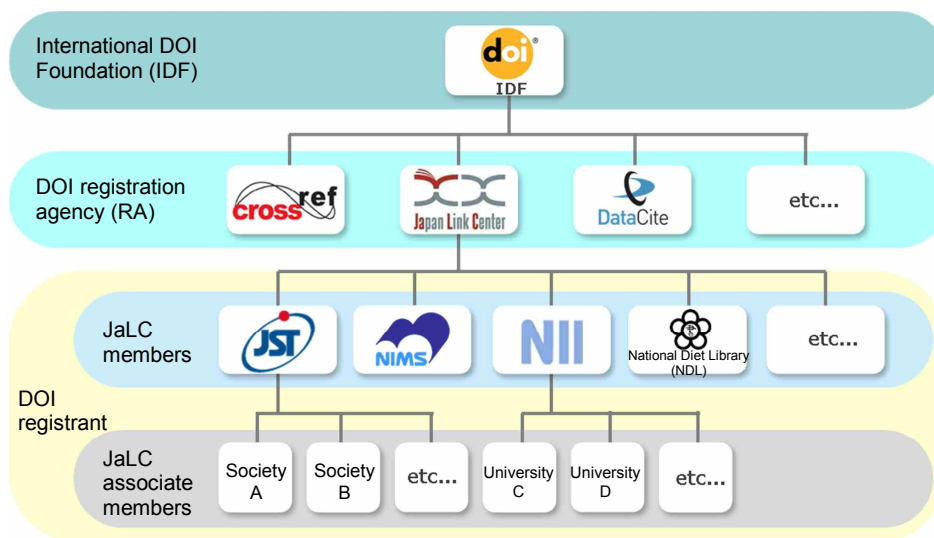


Figure 2-1. Organizational structure of DOI operations

IDF plays a role in creating and managing a database of pairs of DOIs and URLs. The management activities include new registration of DOIs and URLs, acceptance of requests for changes, and maintenance of the function that returns URLs in response to DOI queries.

RAs, which are authorized by IDF, offer services for DOI registration. As of the end of September 2015, nine agencies are registered as RAs around the world. Each RA has their own registration policy including types of content and a range of registrants, and offers services for DOI registration to their registrants. Major RAs include JaLC, CrossRef, which registers DOIs for journal articles on a large scale, and DataCite [1], which registers DOIs for research data ahead of all others.

DOI registrants must become a member of an RA to register DOIs for their own materials. DOI registrants are the publisher or administrator of their materials. Using services that are prepared by RAs, registrants register pairs of DOIs and URLs that are used to access their materials.

JaLC is a member of CrossRef and DataCite, as well as an RA. Members of JaLC can register DOIs assigned by CrossRef or DataCite via JaLC, as well as DOIs directly assigned by JaLC. Figure 2-2 shows the DOI registration flow via JaLC.

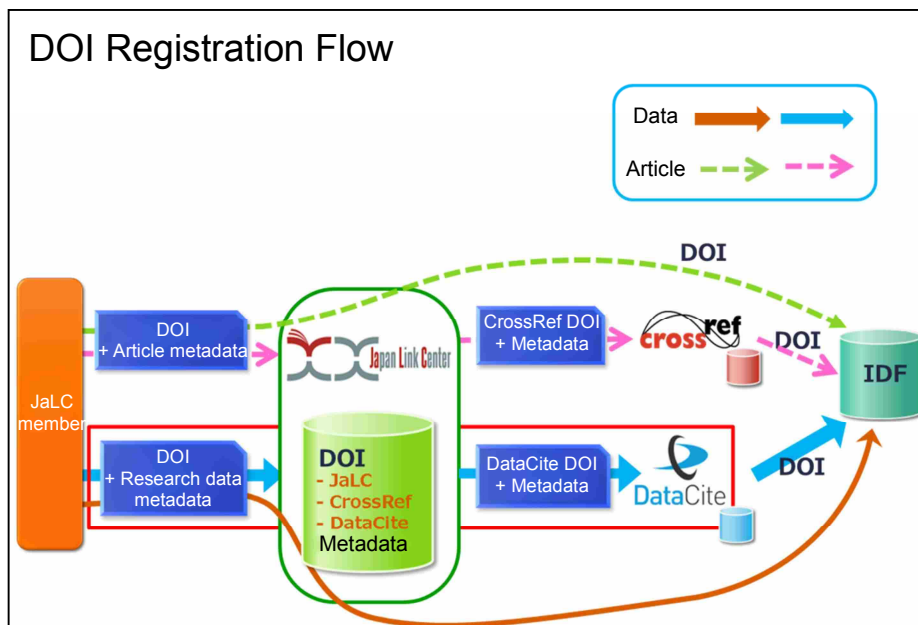


Figure 2-2. DOI Registration Flow via JaLC

Refer to materials [2], [3], and [4] in “5. Bibliography” for details on DOIs and JaLC.

2.2 JaLC metadata

JaLC defines metadata for each type of content and supports journals, journal articles, books and reports, research data and e-learning resources.

Metadata for research data comply with the DataCite schema (DataCite Metadata Schema 3.1). DataCite says that they will provide researchers with methods to obtain the locations of research data and to identify and cite it under the purpose of promoting scientific research, and they define a metadata schema to offer a correct and consistent method for resource citation and search. The mandatory items of the schema consist only of an ID (DOI), the Creator, Title, Publisher, and Publication Year, which are necessary for citation. Minimizing the number of mandatory items, DataCite considers achieving high compatibility with schemas in many fields. However, they define many other recommended and optional items to facilitate data search and citation. [5] [6] [7]

JaLC uses the definition of the DataCite schema with some changes that reflect individual situations, such as coordination with the metadata definition of other content types that are supported by JaLC (such as journal articles), changes in naming rules for tags and methods of supporting multiple languages, and the addition of fund information. The specification of JaLC metadata is

provided in the technical information page on the JaLC website. [8]

3. Guidelines for registering a DOI for research data

3.1 Workflow

3.1.1 Life cycle of data, data creator, administrator, and institution

The DOI registration process starts with creating the content to register followed by preparation of storage and publication, including the creation of metadata, and transmission of DOIs and metadata to an RA. In some cases, a review is conducted to ensure the quality of content during the processes. After registration, the content can be modified or updated with changes to the metadata. It may even be removed. Figures 3-1 to 3-5 show the life cycle of content from data creation to removal and the relation to stakeholders.

The conventional life cycle of literature data such as journal articles and the roles of the person or institution in charge of the process are shown in Figure 3-1 (publisher) and Figure 3-2 (institutional repository managed by a university or research institution). In the former case, the researcher creates articles and the publisher saves and publishes them, registers metadata and a DOI. In the latter case, a researcher creates and saves articles and uploads them to the repository. The library creates and registers metadata, and sometimes registers a DOI as an identifier. Modification and removal of content and metadata must be conducted by the creator.

[Example]

Meanwhile, the processes for research data are complicated. (Figure 3-3) The creator of content (research data) is expected to be the researcher. The creator of the metadata, on the other hand, varies depending on the case, and may be the researcher, or a research assistant, library, data center, or specialist who is in charge of data publication including quality control. JaLC metadata, which are used for DOI registration (metadata for DOI), mainly consist of general items that do not depend on the research area, and metadata for a data repository (domain metadata) contain items specific to the research field and other detailed information.

This document covers processes associated with DOI registration after JaLC metadata creation. The processes from creation of content to its removal, and the creation of domain metadata and its modification are out of the scope of this document. JaLC expects landing pages with links to research data, which is accessed through DOIs registered by JaLC, to contain detailed information about the research data using domain metadata.

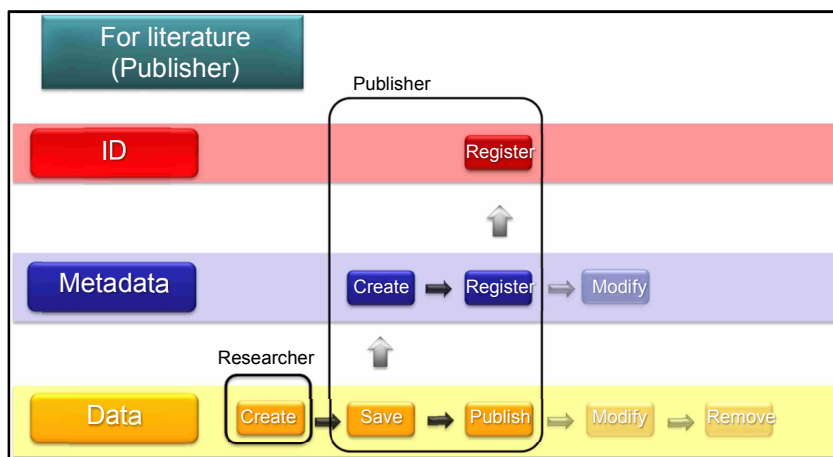


Figure 3-1. Life cycle of literature data (Publisher)

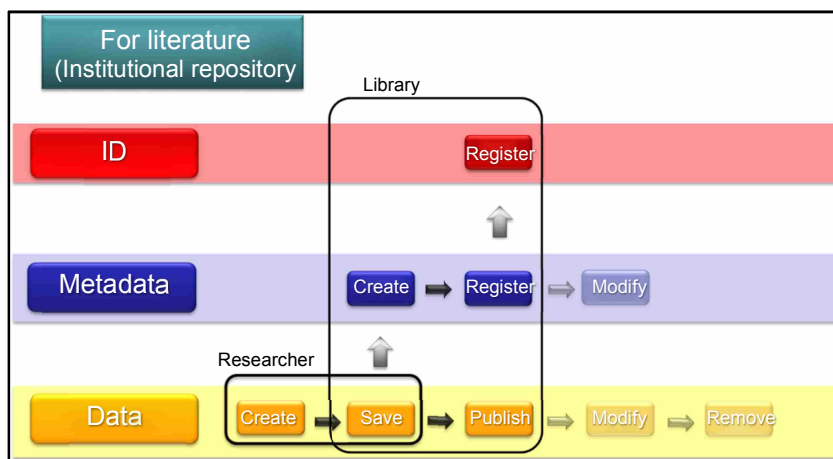


Figure 3-2. Life cycle of literature data (Institutional repository)

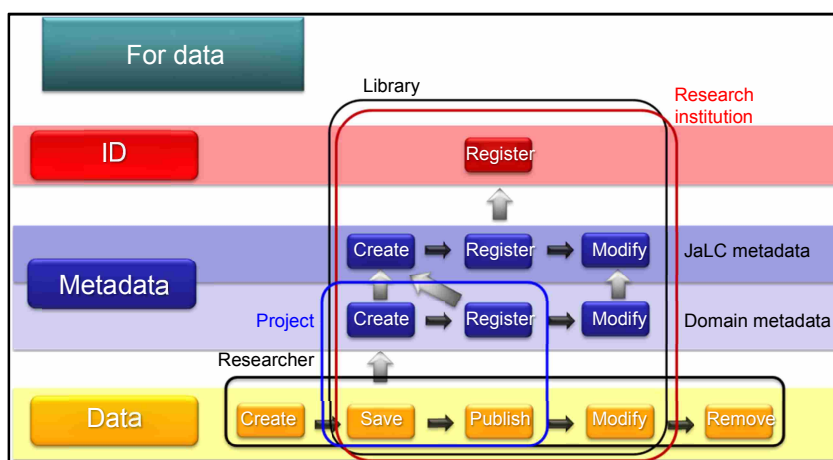


Figure 3-3. Life cycle of research data

3.1.2 Membership of JaLC

In principle, institutions and organizations that do not exist for a limited time are eligible to become members or associate members of JaLC. When the content to be registered is journal articles, there is little need to consider exceptions because the owner is a publisher, university or academic society. Research data may be created by a research project that only exists for a limited time, along with a university or a research institution. Research projects that only exist for a limited time are not eligible to become JaLC members because registered research data must be accessible for a long time. A member institution of the project becomes a member or an associate member of JaLC and registers its DOIs to JaLC in its capacity as a member. When a project that lasts for a long period of time makes an application, the JaLC management committee will judge its eligibility individually, and the project itself may become a member of JaLC.

[Mandatory]

When a project that wants to register DOIs for research data involves multiple institutions, the representative institution must take the overall responsibility. Actual DOI registration, however, can be done by the representative institution, or each institution (in accordance with the rules of the project). In any case, DOI registration must be done by a member or an associate member of JaLC. The representative institution of a project must know which member institution registers DOIs for which research data.

[Mandatory]

Examples of projects that consist of a single institution and multiple institutions are shown below.

- A project in a single institution

[Example]

Figure 3-4 shows an example of a project in a single institution. Project P in Research Institution A will register a new DOI for research data obtained in the project. Since Research Institution A is a member of JaLC and has registered DOIs for their journal articles, it will register DOIs for research data created by Project P in its capacity as a member. The institution applies to JaLC to get a new prefix for research data, which is different from the prefixes for journal articles, and uses the prefix for DOI registration. (Refer to the next section “3.1.3 Principles of prefix assignment” for details of the prefix.) The management systems for journal articles and research data are usually different in an institution. The separate management will lead to smooth operations. After the

project ends, Research Institution A must guarantee that the research data will remain accessible.

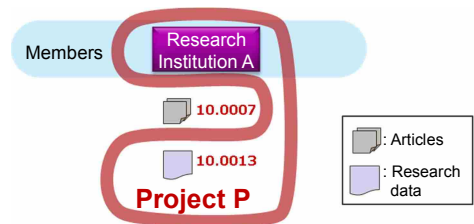


Figure 3-4. A project in a single institution

- A project that consists of members from multiple institutions

[Example]

Figure 3-5 shows an example of Project Q that consists of members from multiple institutions. The representative institution of the project is Research Institution A, which is a member of JaLC. Other members of the project are Society X, University Y and University Z, which are associate members of JaLC under a member of JaLC other than Research Institution A. (Society X is under Institution B, and University Y and Z under Institution C.) When Research Institution A takes over all of the research data obtained during Project Q and conducts DOI registration, only Research Institution A submits an application to JaLC for a prefix, showing that Research Institute A is the representative institution for the project. If the project decides that each institution will conduct DOI registration instead of Research Institution A, Society X, an associate member, will register DOIs via Institution B, a member of JaLC, and Universities Y and Z, associate members, via Institution C, a member of JaLC. These members of JaLC will apply to the JaLC secretariat for a prefix and register the DOI.

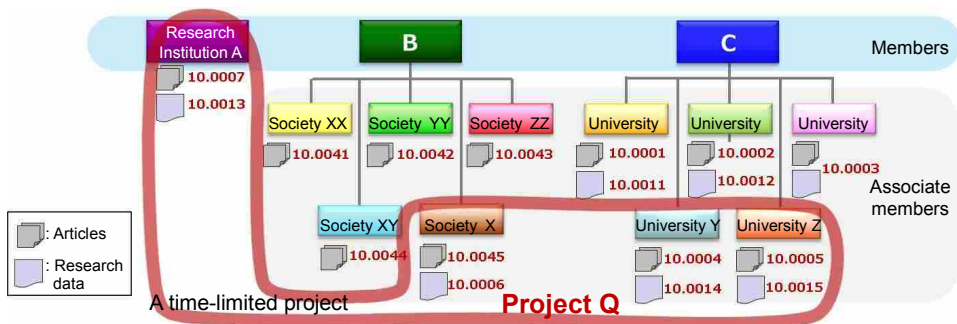


Figure 3-5. A project that consists of members from multiple institutions

Using these examples that show typical patterns of research projects, each

[Recommended]

institution and project is expected to design DOI naming rules including a prefix application policy and establish operational rules that will work for their own situations.

3.1.3 Principles of prefix assignment

The JaLC secretariat assigns DOI prefixes to members of JaLC. In principle, one prefix is assigned to one member. If necessary, multiple prefixes can be assigned.

Research data is created and managed under a variety of situations that are often different from those involved in creating and managing journal articles. It may be created by one laboratory in an institution or a time-limited project that consists of members from multiple institutions, or may be transferred to another institution and be managed there. Considering these situations, it is preferable to use individual prefixes for each administrator of actual data or metadata. For example, each member institution of a project uses an individual prefix, or each group of institutions that creates research data uses an individual prefix if they use a common repository.

[Recommended]

Based on the principle that each administrator uses an individual prefix, a different prefix can be used for each type of content. When an institution, which has registered DOIs for their journal articles using a prefix as a member of JaLC, starts registering DOIs for research data, they can use a different prefix for research data because they have a data management section.

[Example]

When an institution needs a new prefix, it will submit an application with information on the data to be registered to the JaLC secretariat and obtain a prefix. (Associate) members must manage their registered data and metadata properly. Members must report the status of their prefix management to JaLC once a year and return prefixes that are not used and will not be used in the future.

[Mandatory]

3.2 Research data to be registered

3.2.1 Principles of research data to be registered

[Mandatory]

After registration and publication of a DOI, its registrant (administrator) must guarantee that the data will remain accessible. Therefore, the organization must maintain their management system for a long time. (Refer to “3.3 Guarantee of continued access”) Also, research data to be registered must be judged to be necessary and able to be maintained by the organization and have a low chance

of being removed.

3.2.2 Quality of research data

The DOI system guarantees accessibility; it does not guarantee the quality of registered data. However, registrants can write the data quality in the metadata. In some cases, data to be registered is peer-reviewed before DOI registration and the results are written in the metadata.

[Example]

3.2.3 DOI coordination between data repositories

While a referent can be specified by more than one DOI name, it is recommended that each referent have only one DOI name. (DOI Handbook [2], 2.3.5 Uniqueness)

[Recommended]

However, a set of data, especially research data, can be specified by more than one DOI name because several institutions may create data and register the data to their own repositories, which have their own operational rules for DOI registration. Under this situation, it is recommended that data providers communicate with each other and know the overall status of DOI registration. It is also recommended that in cases where the same research data exists with a different DOI, this fact is shown in the metadata as related information (relation_list).

3.3 Guarantee of continued access

3.3.1 Data held by time-limited projects

[Mandatory]

As described in “3.1.2 Membership of JaLC,” DOI registration is conducted by a member of the project who is a member of JaLC. The institution must prepare methods to guarantee continued access.

Unlike with journal articles, research data is sometimes managed by a relatively small group such as a laboratory. Which institution will manage the data becomes an issue, especially in the case of a joint project undertaken with other institutions outside the organizational system of the institution. If data creators or administrators cannot guarantee continued access, outsourcing data management to an appropriate partner institution or data center can guarantee access to the data.

(Examples of data storage after a project ends)

[Example]

- Ex. 1) A representative institution takes over all of the data registered by the project and manages it. Prefixes may be transferred. Actual data

management may be conducted by the member researchers who register the data, an institutional repository, or a library.

Ex. 2) The institutions that conducted DOI registration manage the data.

Ex. 3) If there is no appropriate institution, data management will be outsourced to an external data repository.

Before the project ends, it is useful to establish operational rules such as a system that guarantees continued access after the project is completed.

3.3.2 Removing data after registering a DOI

Even when a data provider ceases publication after registering a DOI, the DOI landing page must be maintained to give continued access by showing the metadata. The landing page must include the fact that the data has been removed and the reason.

[Mandatory]

3.3.3 Data management policy in research activities

Registered data must be properly managed by each member. The methods are not included in these guidelines. See references [9] and [10] shown in Section 5 for details.

[Mandatory]

3.4 Granularity of DOI registration

3.4.1 Basic concept

While the level of granularity for journal articles is easily determined—one DOI is simply assigned to one article—the level of granularity for research data is difficult to determine because it may have different features and be used in various ways. The final decision on granularity is made by the data provider, who must consider that the DOI will be used for a long time.

[Recommended]

3.4.2 Factors in deciding granularity

The following are factors used in deciding granularity of DOI registration.

[Example]

- Citation

Citation linking and measurement of the number of citations is one of the major purposes of DOIs. Registering a DOI for research data makes it easy to cite data and measure the results. Therefore, it is important for data providers to register DOIs with a level of granularity good for citation.

- Data properties

It is recommended that the granularity convey the meaning of the individual

data. However, such granularity varies depending on the data such as observational data, experimental data, or calculation data. For example, while most observational data is not reproducible and data obtained from one observation does not necessarily have an individual meaning, experimental data is expected to be reproducible under the same conditions and data obtained from one experiment may have a clear meaning.

- Accessibility

Needless to say, use of DOIs must guarantee that the content is accessible. Furthermore, it is recommended that the granularity of research data in DOI registration make it easier for users to effectively view and use the data.

- Management

To enable institutions and administrators to guarantee continued access, easy management is important. It is recommended that due consideration is given to easy data management. Some research data is large in scale and is measured and created by many institutions. Registering data that is not managed by an administrator must be avoided.

- Quantity of DOIs

The level of granularity must be adjusted so that the system that registers or resolves DOIs can properly handle them. DOIs must not be assigned to an enormous number of individual measurement records.

3.4.3 Examples of granularity

The following examples give some ideas of DOI registration for research data.

[Example]

(1) Mesospheric wind velocity data observed with MF radar (NICT)

Alaska Project of NICT (CRL)-GI/UAF, Mesospheric wind velocity data (30min. mean) observed with MF radar at Poker Flat, Alaska, doi: 10.17591/55838dbd6c0ad

This dataset is obtained from the radio observation of wind velocity in the upper atmosphere over Alaska since 1998. When the observation device is working, data is added once every thirty minutes or so in semi-real time.

This is a dynamic dataset (a kind of so-called dynamic data) that increases with lapse of time. Generally, a fixed (unchanged) dataset (static data) is suitable for DOI registration. However, such a dataset is difficult to define in

continuous observation or measurement experiments. Therefore, they decided to register one DOI for the whole time-series of data that increases with lapse of time. Observational and measurement data in the geophysics field such as geomagnetic observation data and geomagnetic indices from WDC for Geomagnetism, Kyoto University, and ionospheric observation data from WDC for Ionosphere and Space Weather, NICT, have the same features.

When journal articles cite the dataset, they are expected to give the DOI and access date. Modification records will make it easy to guess the version of evidence data used if the data has been modified after the articles were written.

(2) Spectral Database for Organic Compounds (AIST)

(<http://sdb.sdb.aist.go.jp/>)

Although assigning DOIs to smaller groups of data makes citation easy, it may raise some issues such as difficulties in preparing their landing pages or strange system behavior when data is displayed in a search result. To clarify these issues, they have experimentally registered data 1) using one DOI for the whole site data, 2) per compound, and 3) per spectrum.

The result revealed that if the landing page does not have the function to directly reach the desired data when accessed using the DOI, the URL description becomes very difficult. Especially, when one set of data is displayed using several frames, it cannot be expressed using one URL.

3.4.4 Addition and modification of data after DOI registration

Once data has been registered, it is not expected that the data will change. However, dynamic data can be added or modified later. There are two ways to handle the situation: maintaining the same DOI or registering another DOI.

Using the original DOI after data is added

The following are possible methods to handle data:

- Ex. 1) Use version control. When adding data, or correcting or modifying methods used to obtain or process it, compensation or errors, create a new version, write a version description on the landing page, and provide links to each version of the data.
- Ex. 2) When adding data, or correcting or modifying methods used to obtain or process it, compensation or errors, write the facts and

detailed information as a history record on the landing page.

Ex. 3) When observational data is added at appropriate times, maintain the original DOI as it is and do not perform version control.

It is recommended that these examples are optimized to suit the features of the data and its field.

Registering another DOI after data is added

[Example]

The following are possible methods to handle data:

Ex. 1) Use version control. Write information on each version in the metadata of the original DOI data and the added DOI data.

Ex. 2) When the data is an observational data set that has been obtained over a long time, give another sequential name to the newly obtained data set that contains the data obtained over a certain period of time (e.g. one year) and register DOIs for each data set.

After another DOI is registered, it is recommended that the data specified by the original DOI is stored as it is. However, it may be difficult or even become useless to maintain data that has been entirely modified. When the original data set becomes unable to be accessed for that reason, the landing page must be maintained and the reason must be written on the page.

[Recommended]

3.4.5 Relationship with the quantity of data and the number of files

Granularity of data in DOI registration is not relevant to the quantity of data and the number of files. Although one DOI corresponds to the metadata and the landing page on a one-to-one basis, the actual data to be registered may be a number of files. In that case, one landing page includes one set of metadata and multiple links to the files.

3.4.6 Assigning a suffix

[Example]

A suffix consists of a simple character string to identify the DOI and originally has no meaning nor syntactic structure in itself. Administrators can choose any manageable rules: strings that are randomly generated or strings that show meanings and structure of the data. The latter strings can be easily understood and are intended to facilitate data management and utilization. However, if suffixes include changeable names (i.e. organization name, project name, etc.),

careful consideration is needed because DOIs are used to access data for a long time.

Ex.)

- (prefix)/(sub-organization name).(DB name)-(Identifier in DB)

Ex.: 10.14977/05.tdbs-23732

10.14977/05.gsj-aster-xxxx

- (prefix)/(Acronym of institution).(Serial number)

Ex.: 10.11503/nims.1001

- (prefix)/(random number)

Ex.: 10.17591/55838dbd6c0ad

3.5 DOI landing page

[Mandatory]

Data for DOI registration may be published with or without access restrictions, but the landing page for the DOI must not have access restrictions. Whether the data is accessed with/without restrictions must be indicated on the landing page.

It is recommended that the landing page includes mandatory items of JaLC metadata (DOI, Creator, Title, Publisher, Publication Year) and the following items:

[Recommended]

- Rules for data citation
- Access with/without restrictions
- History of addition, modification and removal
- Data format
- License to use data (Creative Commons license or the license defined by the institution)
- Metadata specific to the research field, etc.

Metadata specific to other fields related to the research field is also useful for researchers and specialists in the field.

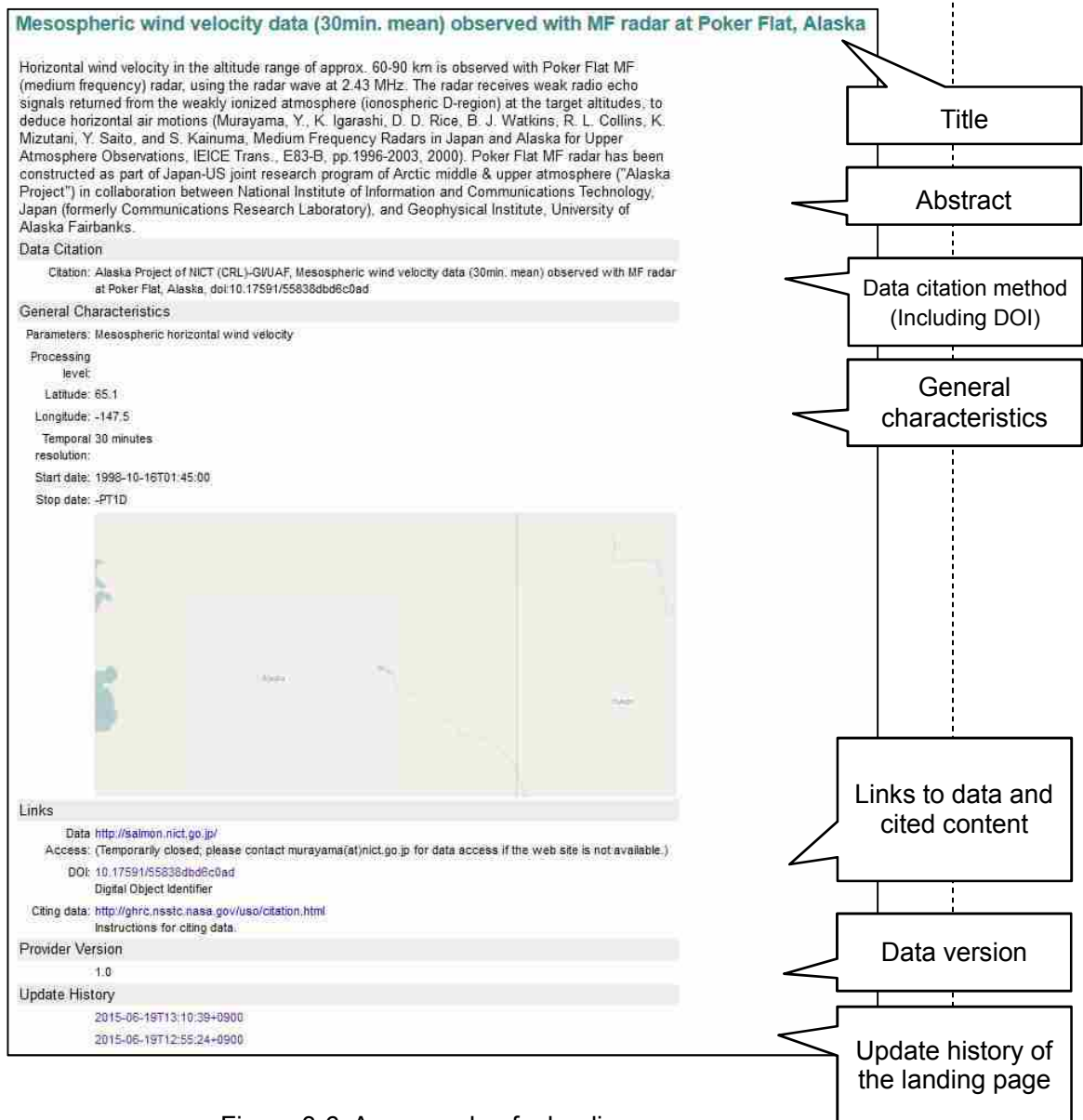


Figure 3-6. An example of a landing page

(Source: doi: 10.17591/55838dbd6c0ad)

3.6 Establishing your own DOI registration policy

[Recommended]

It is recommended that institutions establish their own DOI registration policy in writing before starting to register DOIs in the light of 3.1 to 3.5 above. In the policy, operations and the person responsible in each step of the life cycle of data (data modification timing, application flow, etc.) are defined to ensure that the DOIs maintain accessibility and smooth operations. It is especially useful for each project to establish a data management policy that conforms to the data management policy of the institution before the project ends.

4. Case studies

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

(Features)

- A project undertaken by one institution
- Target: Experimental and observational data on the Arctic and Antarctic
- Systematic DOI registration with JaLC instead of internal data management
- A system that controls data quality has been established.

(Overview of DOI registration)

ADS Metadata Registration System (ADS-AMS) manages metadata and archives data in a unified way. Researchers create metadata and actual data. When they register metadata in the system, the system checks the metadata in accordance with the ADS metadata schema and provides feedback to registrants.

ADS-AMS allows only administrators to register DOIs in accordance with the DOI assignment policy of NIPR. Landing pages for all registered data are created automatically based on their metadata.

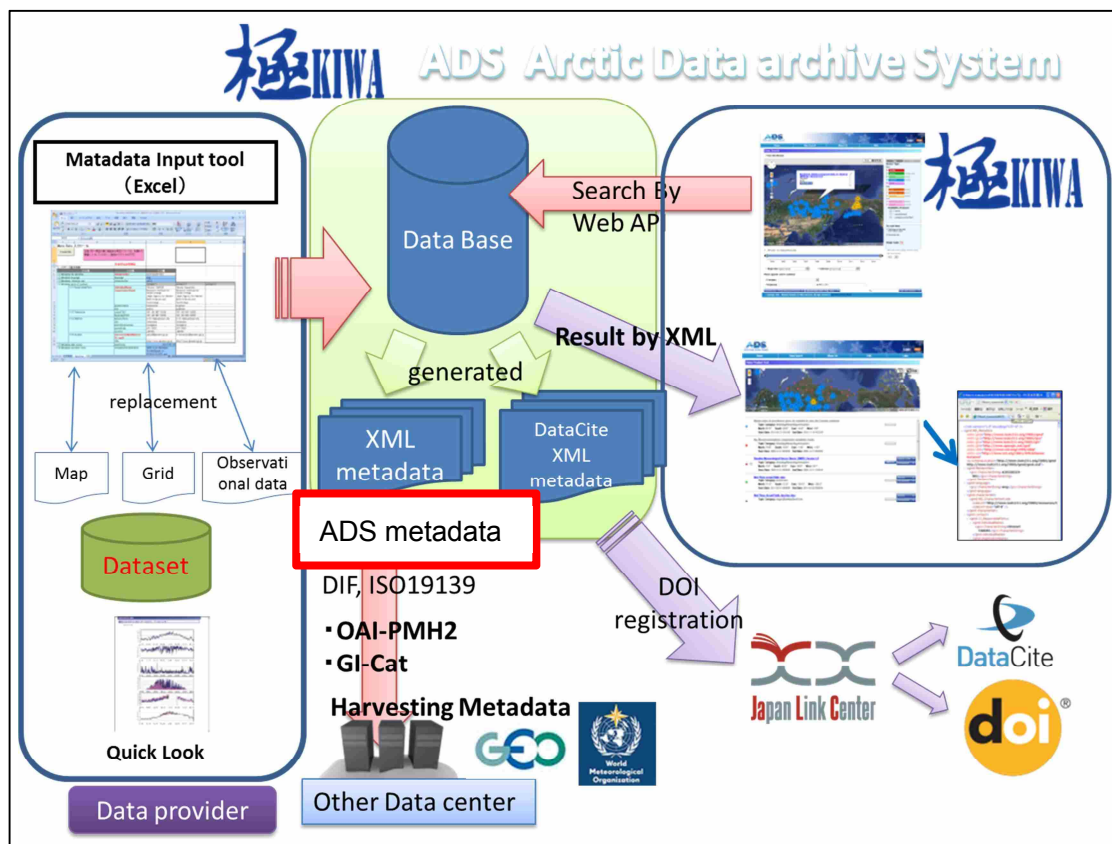


Figure 4-1. Arctic Data Archive System (ADS)

■ Data Integration and Analysis System Program (DIAS-P)

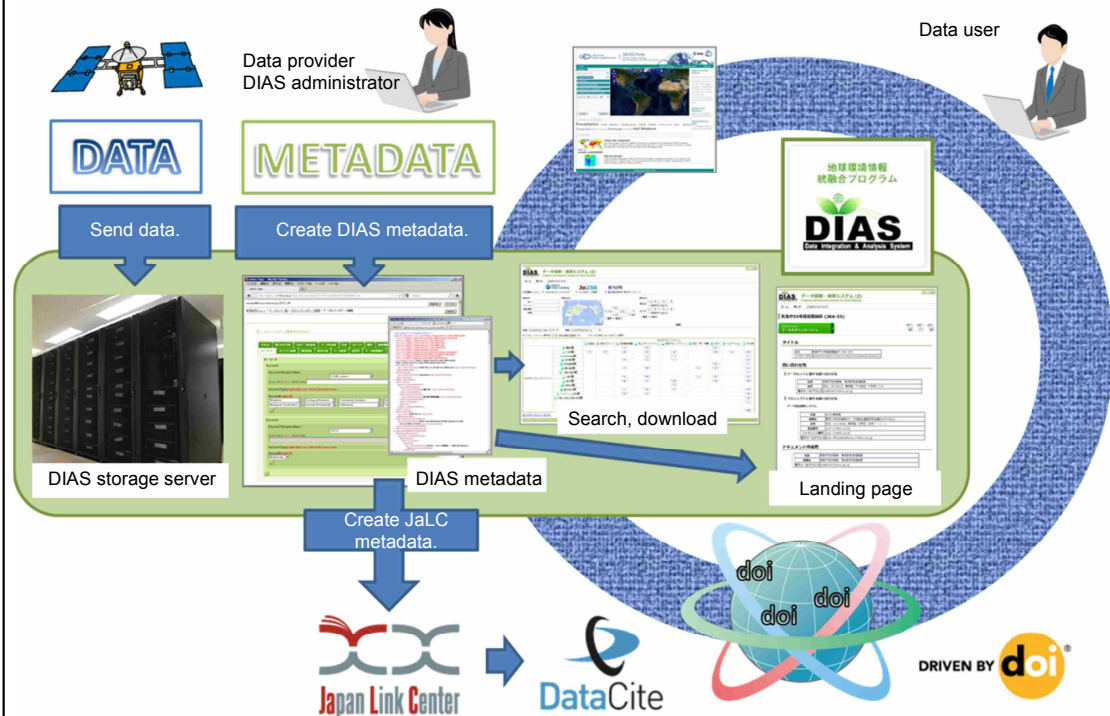
(Features)

- Targets: Data created by DIAS-P itself as the first priority, and data collected or created by other institutions under collaborative projects
- Data provided by other institutions is copyrighted by its creators. DOI assignment and granularity of DOI assignment are determined under an agreement with creators.
- The target field is geoscience (earth observation data, satellite observation data, meteorological prediction models, climate change prediction models and other social data).

(Overview of DOI registration)

- As a practical solution, granularity of data sets to which DOIs are assigned will be the same as for data sets currently provided by DIAS.
- The DOI landing page for each data set is expected to be created from document metadata for the data set created by DIAS.
- JaLC metadata are created from metadata defined by DIAS and registered with JaLC and DataCite during the DIAS data publication process.
- Since the existence of DIAS as a permanent organization is not guaranteed, actual DOI registration is not conducted during the period of this project.

DOI registration with DIAS



Data provider	DIAS administrator	System to use	JaLC/DOI
(1) Request data publication.	Judge whether the data can be published upon request.		
	(2) Discuss file transfer method, data ID, and metadata creator with the data provider	DIAS metadata management system	Choose DOIs to be assigned to data.
(3) Send files to DIAS.	Store the files.	DIAS storage server	
	(4) Ask the data provider to create metadata.		
(5) Input metadata.	Check the metadata and progress.	DIAS metadata management system	
	(6) Allocate the files to the DIAS public server.	Storage server that can be accessed from outside of DIAS	
	(7) Check files to be downloaded and register the files in the system that uses the files.	DIAS data download system Search and Discovery System for DIAS Datasets Applications developed by DIAS for each data	
(8) Check terms of data use and citation and ensure that all understand DIAS terms of use and citation.		DIAS metadata management system (JaLC registration metadata output function) (Landing page creation function) Search and Discovery System for DIAS Datasets	Register DOI, JaLC registration metadata, and landing page address to JaLC
	(9) Set individual access privileges to the data for each application (no authentication, agreement with terms of use, application for use).	DIAS data download system Applications developed by DIAS for each data (Landing page edit function)	
(10) Publicize data. Users run applications to use and download data.			
(11) Users can resolve data using DOIs, which are available for citation.			

Issues: Specification (interpretation) changes to integrate new items into DIAS metadata (ISO19139)

Figure 4-2. DOI registration with DIAS

■ National Institute for Materials Science (NIMS)

(Features)

- Target: Images obtained with electron microscopes in materials science
- Collaboration between the existing data archive system and DOI registration system

(Overview of DOI registration)

A separate DOI management system has been established to provide a DOI registration function via “NIMS eSciDoc,” an institutional repository and self-archive system without a DOI registration function.

Registration and publication of research data is done by researchers, and DOI registration is done by the Scientific Information Office.

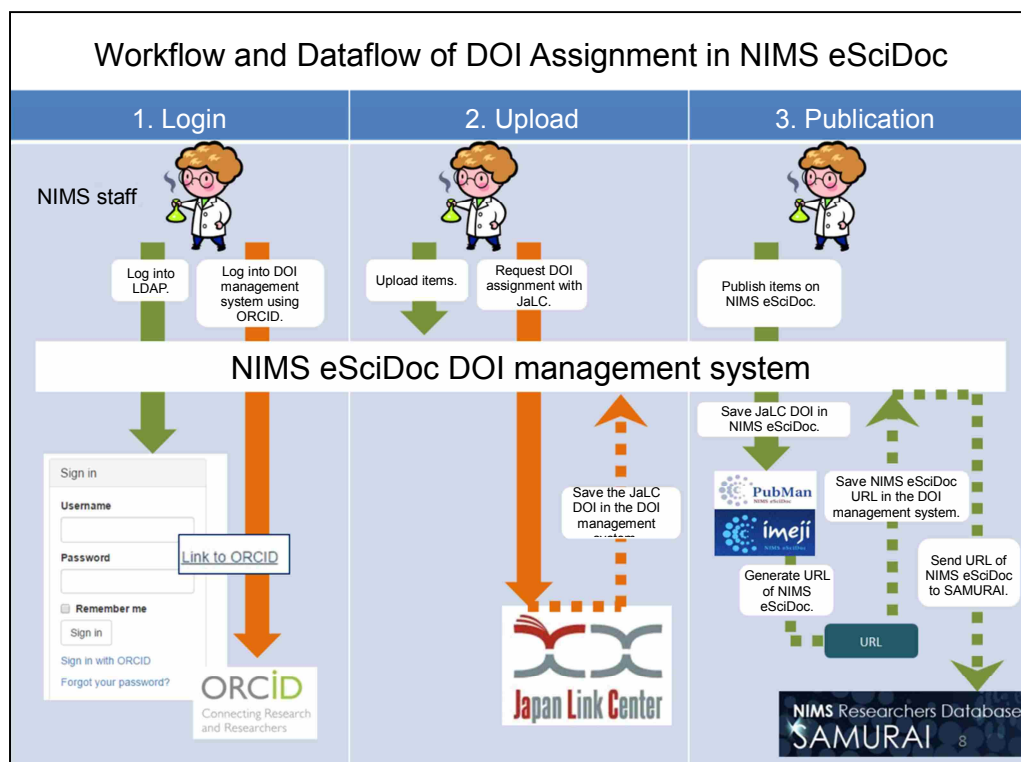


Figure 4-3. Workflow and Dataflow of DOI Assignment in NIMS eSciDoc

Note: Actual DOI registration is under consideration.

- RIKEN Brain Science Institute (BSI)
Neuroinformatics Japan Center (NIJC)

(Features)

- Target: Data contained in databases for neuroscience
- DOI registration support function using the existing database building system is under consideration.
- Peer review of research data will be conducted.

(Overview of DOI registration)

Using data registration, review, and management functions of XoonNips (a Web database building system), building a system to support processes until assignment of DOIs to research data for each platform (PF) is under consideration. The flow of DOI registration includes provision of metadata from operators of each PF, review of data to check conformity with the DOI registration policy, and management of registered DOIs.

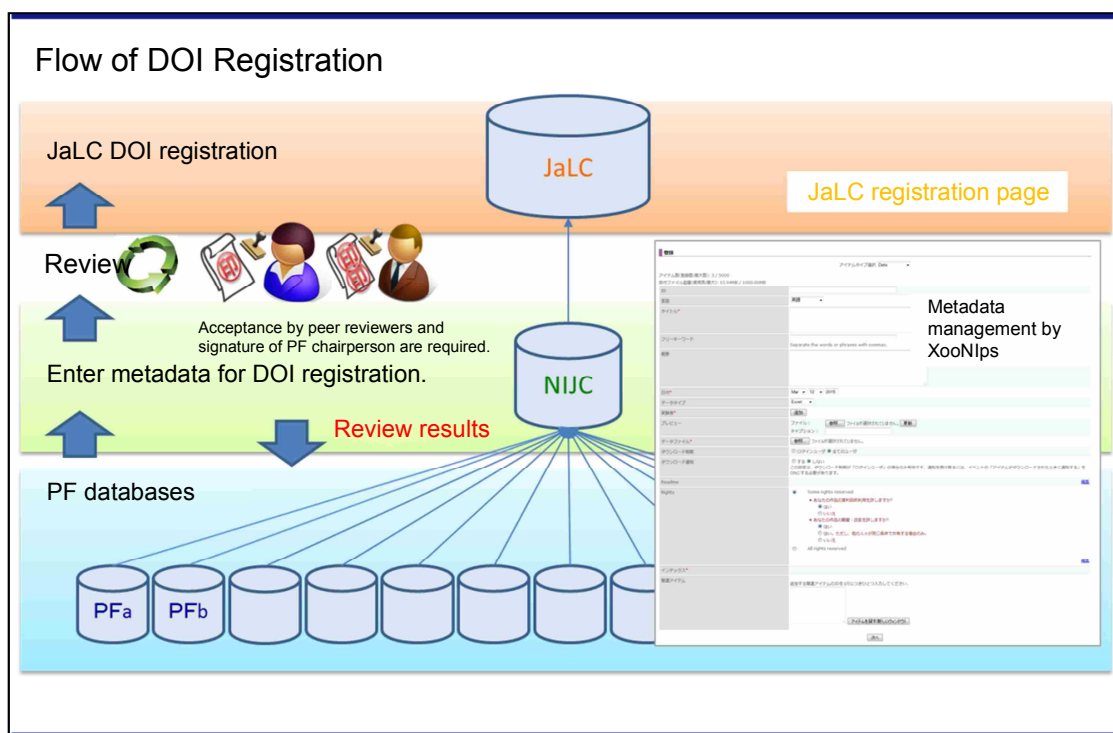


Figure 4-4. Flow of DOI Registration in NIJC

Note: Actual DOI registration is under consideration.

5. Bibliography

- [1] DataCite, "DataCite," [Online]. Available: <http://www.datacite.org/>. [Accessed: May 25, 2015].
- [2] ©International DOI Foundation, "DOI®Handbook," 17 March 2014. [Online]. Available: <http://www.doi.org/hb.html>. [Accessed: June 28, 2015].
- [3] Japan Link Center. [Online]. Available: <https://japanlinkcenter.org/top/index.html>. [Accessed: May 25, 2015].
- [4] Japan Link Center, "What is the Japan Link Center? – History and principles," July 28, 2014, [in Japanese]. [Online]. Available: http://doi.org/10.11502/jalc_policy. [Accessed: May 25, 2015].
- [5] DataCite, "DataCite Metadata Schema for the Publication and Citation of Research Data," June 2015. [Online]. Available: <http://doi.org/10.5438/0010>. [Accessed: September 10, 2015].
- [6] DataCite, "DataCite Metadata Schema Repository," [Online]. Available: <http://schema.datacite.org/>. [Accessed: May 25, 2015].
- [7] J. Starr, "isCitedBy: A Metadata Scheme for DataCite," January/February 2011. [Online]. Available: <http://doi.org/10.1045/january2011-starr>. [Accessed: September 10, 2015].
- [8] Japan Link Center, "JaLC Technical Information," [in Japanese]. [Online]. Available: https://japanlinkcenter.org/top/admission/index.html#member_technical. [Accessed: September 10, 2015].
- [9] MEXT, Government of Japan, "Guidelines for Responding to Misconduct in Research," August 2014, [in Japanese]. [Online]. Available: http://www.mext.go.jp/a_menu/jinzai/fusei/___icsFiles/afieldfile/2015/07/13/1359618_01.pdf [Accessed: May 22, 2015].
- [10] Cabinet Office, Government of Japan, "The Expert Panel on Open Science, based on Global Perspectives," 2015, [in Japanese]. [Online]. Available: <http://www8.cao.go.jp/cstp/sonota/openscience/>. [Accessed: May 11, 2015].

Glossary of Terms

CrossRef

A leading RA in the US. Mainly handles articles for DOI name registration.
(<http://www.crossref.org/>)

DataCite

An RA that handles research data. The consortium, which was founded in 2009 by European universities and libraries, is operated by the German National Library of Science and Technology (TIB). (<http://www.datacite.org/>)

DOI

Acronym for “Digital Object Identifier”. Refer to the DOI Handbook.

Identifiers that are permanently assigned to materials on the Internet.

A typical DOI name is a character string such as “10.1246/nikkashi1898.1.1.” The prefix “10.1246” is a directory indicator that is assigned by the International DOI Foundation. The suffix “nikkashi1898.1.1” is an identifier that is assigned by the entity (in this case, the Chemical Society of Japan) that registers the DOI name. When you resolve a DOI name using your browser, the DOI name is appended to URL “<http://doi.org/>”, i.e. <http://doi.org/10.1241/johokanri.56.881>.

DOI prefix

A directory indicator that is assigned by IDF. Refer to “DOI.”

DOI suffix

A directory indicator that is assigned by a registrant. Refer to “DOI.”

DOI registration

Assigning a DOI name to content and registering (depositing) it with URL and metadata with an RA.

DOI resolution, DOI resolver

DOI resolution is a process to obtain the URL of the Web page on which the content or metadata are present by submitting a DOI name. A DOI resolver is a network service that resolves DOI names.

IDF

Acronym for “the International DOI Foundation.” (<http://www.doi.org/>)

JaLC system

An Internet service that manages domestic electronic documents using DOIs.

JaLC is the acronym for “Japan Link Center.”

prefix

Same as DOI prefix in the context of this document. Refer to DOI prefix.

assignment and return of a prefix

IDF assigns prefixes to each RA, which assigns the prefix to an organization or a project that has requested the registration of a DOI name.

RA

Acronym for “(IDF) Registration Agency,” which offers services for DOI registration.

registrant

An organization that requests and receives the registration of DOI names for their journal articles, books, reports and other electronic documents.

suffix

Same as DOI suffix in the context of this document. Refer to DOI suffix.

deposit

To deposit, which means to put something where it will be safe, means to register electronic data in the context of this document.

For example, a registrant deposits metadata and main body converted into the XML format with the JaLC system.

multiple resolution

Generally, a DOI name links to a single Web page and a resolver returns the URL value.

However, when a registrant wants to link a DOI name to multiple URLs (e.g. when a journal article is available from multiple web sites), the DOI name can link to an intermediate page that has all the different associated URLs using multiple resolution.

metadata

The data that provide an overview of information for information retrieval systems is called metadata. In the library and information science field, it is sometimes called bibliographic information. For example, metadata of a document usually includes the author, title, publication date and keywords.

metadata schema

A data structure for metadata. In the information retrieval field, metadata schemas are usually defined using XML.

landing page

A Web page that links to a DOI name and is displayed when the DOI name is resolved.

repository

A service that preserves and provides journal articles, research data and other contents.

Examples are: institutional repositories, which preserve and provide digital materials created by libraries and institutions, and disciplinary repositories, which preserve and provide works and data associated with these works of scholars in a particular subject area.

Japan Link Center

Experimental Project to Register DOIs for Research Data

Members of the project:

Japan Science and Technology Agency (JST)

National Institute of Polar Research (NIPR)

Cyber Science Infrastructure Development Department, National Institute of Informatics (NII)

Digital Content and Media Sciences Research Division, NII

(Members of the Data Integration and Analysis System (DIAS) Project)

NII

Earth Observation Data Integration & Fusion Research Initiative, University of Tokyo

Japan Agency for Marine-Earth Science and Technology

Graduate School of Informatics, Kyoto University

Center for Global Environmental Research, National Institute for Environmental Studies (NIES)

National Institute of Advanced Industrial Science and Technology (AIST)

National Institute of Information and Communications Technology (NICT)

(Members of WDC)

Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science,
Kyoto University

WDC for Ionosphere and Space Weather, NICT

Data Center for Aurora, NIPR

Center for Science-satellite Operation and Data Archive, Japan Aerospace Exploration Agency
(JAXA)

Chiba University Libraries

National Institute for Materials Science (NIMS)

Neuroinformatics Japan Center (NIJC), RIKEN Brain Science Institute (BSI)

(Not in particular order)

Period:

October 2014 - October 2015