

# Technological and community-based approaches to metadata infrastructure for research data

October 14th, 2025

Satoko Fujisawa Japan Link Center (JaLC) Dept. for Information Infrastructure, Japan Science and Technology Agency (JST)



# Overview of JaLC

-Japan's DOI Registration Agency-



# JaLC is Japan's DOI Registration Agency



- JaLC is the <u>DOI RA of Japan</u> accredited by DOI Foundation in 2012
- JaLC registers DOIs for <u>academic content</u> produced <u>in Japan or in Japanese</u>
- JaLC DOIs accounts for 3% of total DOIs in the world (as of 2025 April)

Content-oriented RAs



**Scholarly Publication** 



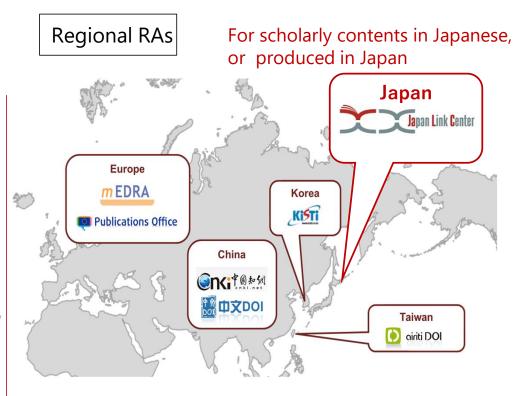
HUMAN Research Data and more

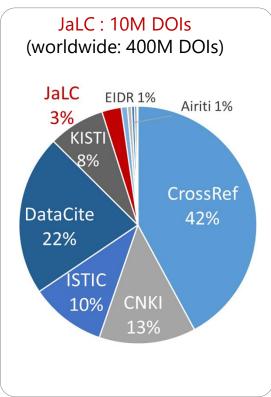
Talent ID



Content ID Video Service ID **bsi.**identify

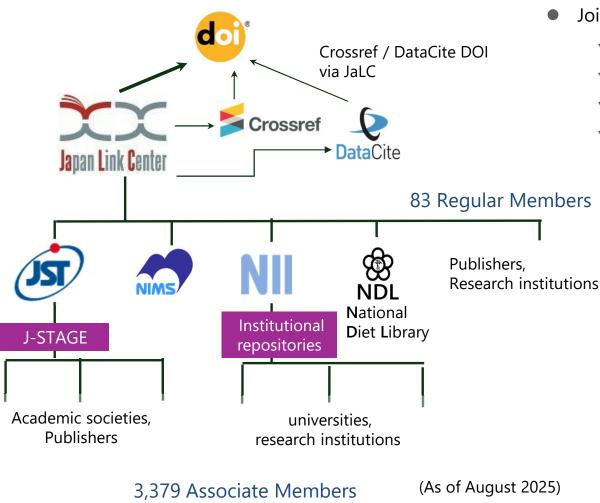
Construction products





### Structural Overview of JaLC



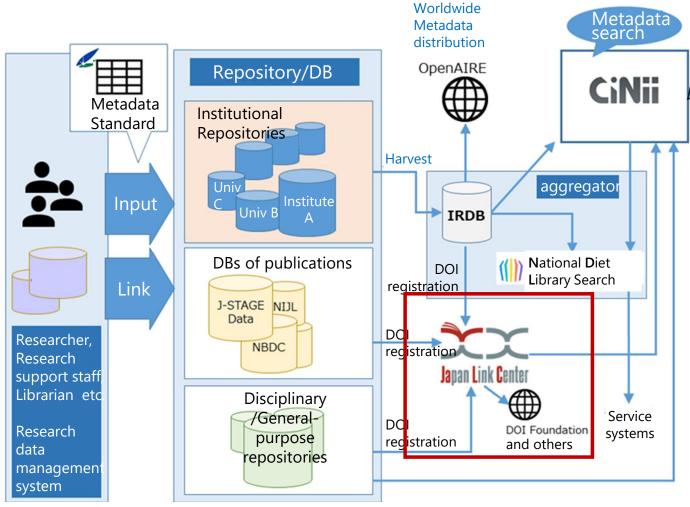


- Jointly managed by 4 governmental institutions
  - ✓ Japan Science and Technology Agency (JST)
  - √ National Institute of Material Science (NIMS)
  - √ National Institute of Information (NII)
  - √ National Diet Library (NDL)
    - JaLC member institutions
    - Some members are
      - √ National research institutions
      - ✓ Publishers of English journal
    - · Most of Associate members are
      - √ Local universities
      - ✓ Small academic societies and publishers
      - → A significant portion of JaLC's content and metadata are in Japanese.

JaLC is optimized for domestic needs

### JaLC is the infrastructure for metadata distribution in Japan





Academic information search service

Provides cross-search functionality for various research-related information, such as bibliographic info, research data, etc.

**CiNii** provides information to other services like researchmap (the comprehensive researchers' catalogue)

Input to JaLC

DOI registration via platforms

- ✓ NII IRDB for Institutional repositories
- ✓ **J-STAGE** for Japanese publishers and other repositories
- Output from JaLC
   Provide metadata to other RA and applications
- ✓ CiNii Research and other services
- ✓ DataCite, Crossref, DOIF

### DOI registration for Research Data in JaLC



- Research Data is one of the five content types in JaLC.
- DOI registration for Research Data has been increasing over the years.

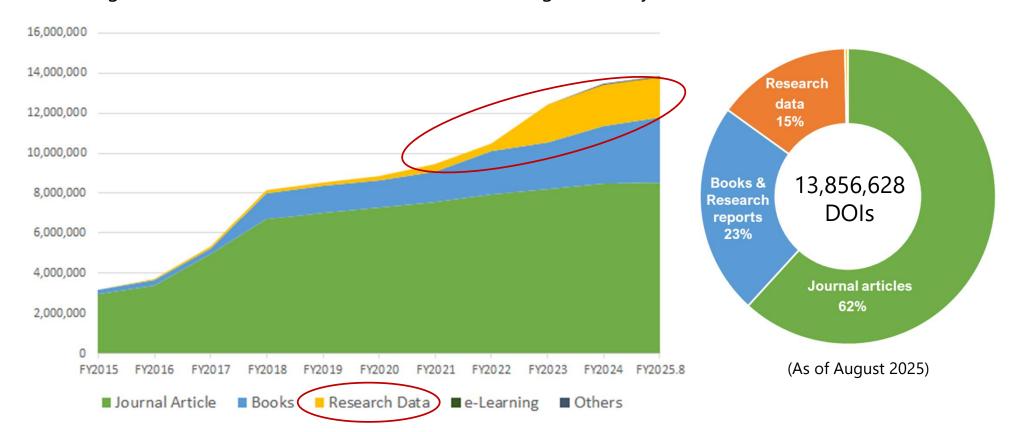


Fig. Total number of DOIs registered in JaLC

<sup>\*</sup> It includes Crossref DOI and DataCite DOI registered via JaLC.

# Compatible metadata schema



- Metadata schema is designed for each content type.
- In principle, metadata for Research Data is defined conforming to DataCite definition.

Contents type	Additional Metadata	External Deposit	
Journal Article	journal name, ISSN, volume, number, title,page, etc.	CrossRef	
Book	series title, chapter, ISBN, etc.	CrossRef	
Research Data	-, 9 , 9 ,		
e-Learning	Learning Resource type, rights, etc.		
Other	(Basic metadata only)		

#### Basic metadata:

DOI, URL, title, author information, affiliation, researcher id, publication date, publisher, edition, related contents, funder

#### JaLC follows DataCite schima update

Currently supporting schema 4.6

"Research Data" is becoming to cover more general resource

according to DataCite update.

- JaLC will cover Preprint as "Research Data"
- JaLC will cover Grants in manner of DataCite ( as "Research Data" )
- \* JaLC's <u>"Research Data"</u> doesn't cover the resource type equivalent to Journal Article or Book

## Metadata enrichment in alignment with Japan's Policy



•JaLC alignes Japan's policy as the infrastructure for metadata distribution in Japan.

National Policy on Promoting Open Access to Publicly Funded Scholarly Publications and Scientific Data. approved by the Cabinet in February 2024

- This policy mandates immediate open access to peer-reviewed publications and <u>research data</u> <u>underlying publications</u> resulting from publicly funded researches from FY2025.
- The Common metadata elements (for research data) has defined in Metadata Manual for
- the Management and Utilization of Publicly Funded Research Data (Version 1.0, July 2024).

	Elements	Mandatory/ Optional	
1	Funder	Mandatory*	
2	Funding Stream Code in Japan Grant Number	Optional	
	Program Name	Optional	
3	Japan Grant Number	Mandatory*	
	Project Name	Mandatory*	
4	Data No.	Mandatory	
5	Title	Mandatory	
6	Date (Issued/Updated)	Mandatory	
7	Description	Mandatory	
8	Research Field	Mandatory*	
9	Data Type	Mandatory	
10	File Size	Optional	

11	Data Utilization and Provision Policy	Mandatory	
	Access Rights	Mandatory	
	Available Date	Mandatory	
12	Repository Information	Mandatory	
12	Repository URL/DOI Link	Optional	
4.2	Creator Name	Optional	
13	Creator Name Identifier (e-Rad)	Optional	
14	Hosting Institution	Mandatory	
	Hosting Institution Code	Optional	
	Data Manager	Mandatory	
	Data Manager Identifier (e-Rad)	Optional	
	Contact of Data Manager	Mandatory	
15	Remarks	Optional	

The content type "Research Data" in JaLC covers these metadata elements.

Source:

https://www8.cao.go.jp/cstp/metadatainstructions\_en.pdf

## Community Support by JaLC



 Dialogue and Co-Creation Forum (JaLC member meeting)



JaLC Seminar for beginners (Webinar)

- ✓ What's PID?
- ✓ Significance and function of DOI
- ✓ About JaLC
- ✓ Q&A



RDUF Symposium



- JaLC NEWS (monthly mail magazine)
  - ✓ Administrative Notices (only for members)
  - ✓ Service Information
  - ✓ Event News
  - ✓ Pick up FAQs



### Community Support: Research Data Utilization Forum



 RDUF is a community that facilitates information sharing and discussions to address issues related to the open sharing and reuse of research data. (since 2016)



• RDUF has focused primarily on the promotion of persistent identifiers for research data.

### Member

190 Individuals



8 Institutes

Variety of Scientific Domains

- Informatics
- **Bio Science**
- Earth & Environmental Science
- Material Science
- **Humanities & Social Science**

#### Variety of Working Positions

- Researchers
- Librarians
- **Publishers**
- IT Developers & Technicians
- Fund & Policy Makers



Gather domain/position specific cases

#### Subcommittee

a short-term working group to tackle an emerging issues, especially among Japanese academia and research communities.



Generalized their domain-specific practices

**Deliverables** 

Guidelines, Reports, infographics... are open to the public



Feedback and outreach/ Facilitate common knowledge

relationships

Academic/Social research organizations, policymakers, universities, academic libraries, and industries.

#### **Example RDUF Outputs & Recommendations**

■ Guidelines for specifying conditions of use in research data publishing



(Research Data License Subcommittee)

DOI: https://doi.org/10.11502/rduf license guideline en

Guideline for Development and Operation of Research Data Repository

(Subcmt.: Japan DAta Repository Network)

Japanese Translation of PIDINST Whit Paper, ePIC cookbook, DataCite cookbook

(Subcmt.: PIDs for Research Materials and **Experimental Equipment)** 

### "Guidelines for the DOI Registration for Research Data" (June 2024)



- The Guidelines are a revision of the version published in 2015.
- The guidelines offer comprehensive and accessible explanations regarding to DOI registration
  - ✓ the foundational information regarding JaLC and DOIs
  - ✓ the principles underlying the assignment of DOIs to research data
  - ✓ procedural aspects in DOI registration.
  - ✓ case studies on DOI registrations in research institutions
- Much of the content can be common with DOIs beyond JaLC and will be useful to many readers who are considering registering DOIs for research data.

https://doi.org/10.11502/rd\_guideline\_en

#### Intended readers for each chapter

	Main intended readers				
Chapter title	(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	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	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		•	•		



# Advancing Global Metadata Sharing

-DOI Registration in DataCite through JaLC-



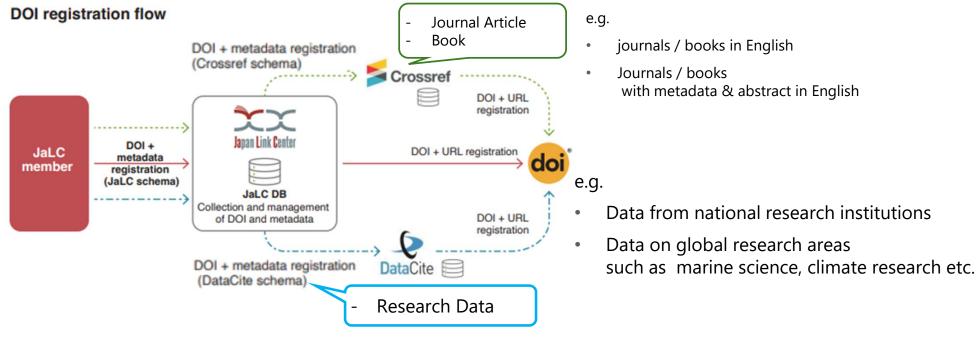
# For enhancing international metadata distribution



#### Support for DOI Registration to Crossref / DataCite:

JaLC supports members in registering DOIs with Crossref or DataCite, if desired.

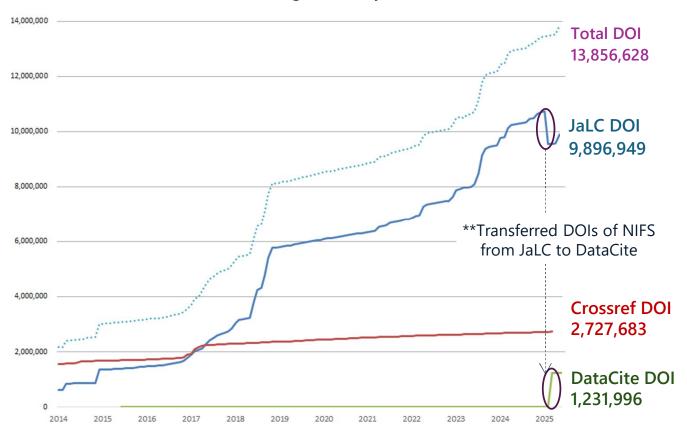
- → JaLC members can register Crossref DOIs or DataCite DOIs within their JaLC membership contract
  - All metadata is stored in JaLC and is distributed to Japan's information networks.

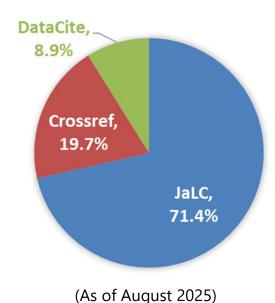


# Registered DOIs in JaLC by RA



- Over 70% of DOIs are registered as JaLC DOI.
- About 9% of DOIs are registered to DataCite through JaLC
- Most of DataCite DOIs are registered by National Institute for Fusion Science (NIFS).

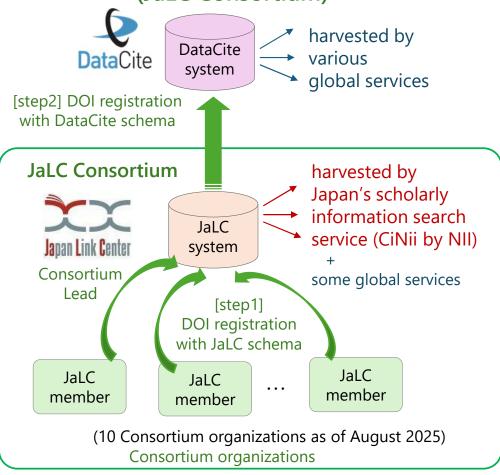




## Metadata schema translation from JaLC to DataCite



# Registration to DataCite through JaLC (JaLC Consortium)



#### Merit

- Members to follow only the JaLC schema.
- Technical support by JaLC
- All the metadata is also stored in JaLC.
  - → findable in Japan's information sevices.

#### Demerit

- Not fully covered in the DataCite shcema
- Some time lag for the translation to catch up the latest DataCite schema.
- Only Research Data can be registered to DataCite (JaLC doesn't cover the resource type equivalent to Journal Article or Book)
- Some advanced users might be frustrated with the schema not being fully covered in the DataCite schema and with the time lag.

## **Current and Ongoing Priorities**



### Toward maximizing the FAIR Principles of JaLC Content

- Provide metadata registration functionalities considering both registration and utilization aspects
  - Continuous updates of metadata schemas
    - → to ensure compatibility both with other PID services and to respond to national policies
  - ➤ Metadata registration support ... manuals, registration interfaces, guidelines
- Expansion of API capabilities to promote JaLC metadata harvest
  - ➤ All the content of JaLC is harvested in OpenCitations Index.
  - ➤ JaLC is also in the process of being harvested by multiple global aggregators, including OpenAIRE.