

# DataCite DOI

## 事例報告

岡山大学附属図書館

大園 隼彦

# 本日の内容

- DataCite DOIの申請・登録
- JaLC DOIとの違い
- (なぜ研究Dataを公開するのか)
- (なぜDOIを付与するのか)

# DOIのおさらい（と岡山大学の事例）

<https://ousar.lib.okayama-u.ac.jp/63011>

||

<https://doi.org/10.48696/63011>

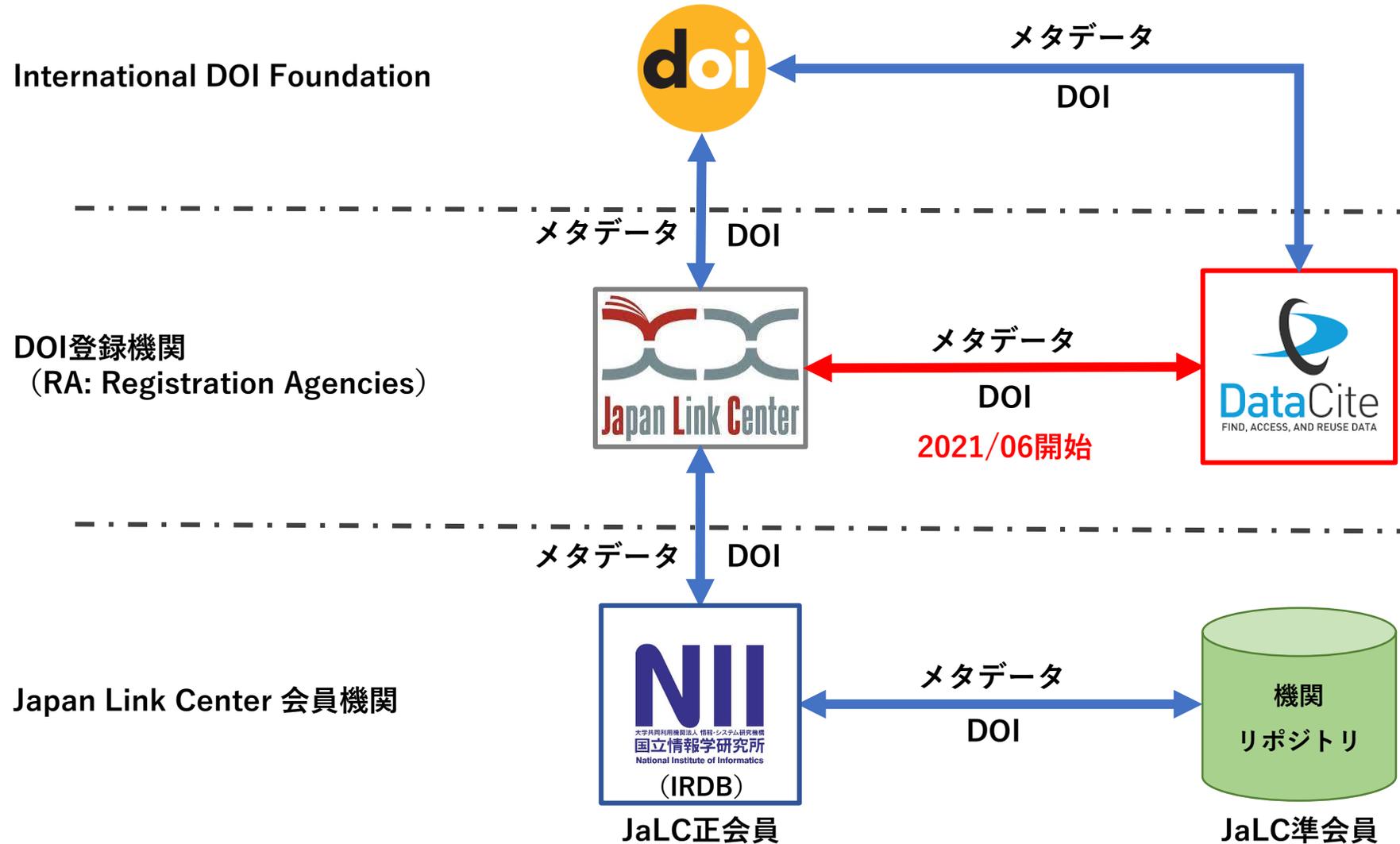
The screenshot shows the OUSAR website interface. The main content area displays the title '日本の雑草 種子画像データベース' and a highlighted DOI: <http://doi.org/10.48696/63011>. Below this, a table provides details for the record:

ID	63011
タイトル	Naturalized plants in Japan : Seed-Image database
JaLCDDOI	10.48696/63011
フルテキストURL	Seed_DB.xlsx 323 KB
著者	山下 純 岡山大学 資源植物科学研究所 <a href="#">Kaken ID</a> <a href="#">publons</a> <a href="#">researchmap</a>
抄録	On this database, there are Seed-images of naturalized alien plants in Japan.
抄録（別表記）	日本の雑草の種子画像を公開しています。帰化植物種子画像データベースと日本の雑草種子画像データベースを合併しました。

○ 学术论文、図書、研究データ等のコンテンツに付与される国際的な識別子

○（機関リポジトリが）DOI - URLの対応とコンテンツの継続的な管理を行うことで、恒久的なアクセスを実現

# JaLCとDataCite



JaLCでもDataCiteでも、研究データにDOIを付与することは可能



何が違うのか？

# DataCite DOIの申請

- JaLC準会員は、DataCite DOIを登録可能に（2021.6）
  - ※ただし、JAIRO Cloud利用機関は全機関のWEKO 3 移行後
- 登録費用なし
- 以下から申請（有効まで1ヶ月程度）

コミュニティサポート by JPCOAR and NII

<https://community.nii.ac.jp>

DataCite DOIの登録開始について

<https://support.irdb.nii.ac.jp/ja/news/20210601>

# DataCite DOIの登録（メタデータ対応）

- 提供するメタデータ

○IRDBデータ提供機関のための DOI管理・メタデータ入力ガイドライン

junii2編. <https://jpcoar.repo.nii.ac.jp/records/215>

JPCOARスキーマ編. <https://jpcoar.repo.nii.ac.jp/records/160>

→ ハーベスト後、有効になるまで1週間程度

参考：DataCite Metadata Schema v.4.x

<https://schema.datacite.org/>

# メタデータ : junii2

junii2ガイドライン メタデータ項目					
通番	項目	要素名 (element)	必須	繰返し可否	形式等
1	タイトル	title	必須	1	・lang属性を設定することが必須。
3	作成者	creator	必須	1以上	・英語名(lang属性"en")での記入を推奨。
4	著者キーワード	subject	あれば必須	0以上	
15	内容記述	description		0以上	
16	公開者	publisher	必須	1	・英語名(lang属性"en")での記入を推奨。 ・複数ある場合、JaLCシステムには、初出のみが取り込まれる。
17	寄与者	contributor	あれば必須	0以上	・英語名(lang属性"en")での記入を推奨。
18	日付	date	あれば必須	0以上	
20	国立情報学研究所 メタデータ主題語彙集 (資源タイプ)	NIItype	必須	1	NIItypeが以下のいずれかの場合、JaLCシステム上、「研究データ」というコンテンツ種別で取り込まれる。 ・Data or Dataset ・Software
21	フォーマット	format		0以上	
23	資源識別子URI (アイテム表示画面へのリンク)	URI	必須	1	
24	本文フルテキストへのリンク	fullTextURL	必須	1以上	
25	JaLCDOI	selfDOI	必須	1	・ra属性に"DataCite"を設定する。 ・半角の"a-z", "A-Z", "0-9", "-", "_", "/"のみ使用可能。 ・プレフィックスは自機関に割り当てられたものに限り使用できる。
34	刊行年月日	dataofissued	あれば必須	0~1	・複数ある場合、JaLCシステムには、初出のみが取り込まれる。 ・ガイドラインに準拠しない場合及び空欄の場合は、JaLCシステムにて発行年に「9999」と設定される。
36	言語	language		0~1	・複数ある場合、JaLCシステムには、初出のみが取り込まれる。 ・空欄の場合、または、対応する言語コードがJaLCシステムにない場合は、JaLCシステムにて「unk」と設定される。
39	DOI	doi	あれば必須	0以上	
42	異版である	isVersionOf	あれば必須	0以上	
43	異版あり	hasVersionOf	あれば必須	0以上	
44	置換される	isReplacedBy	あれば必須	0以上	
45	置換する	replaces	あれば必須	0以上	
46	要件とされる	isRequiredBy	あれば必須	0以上	
47	要件とする	requires	あれば必須	0以上	
48	部分である	isPartOf	あれば必須	0以上	
49	部分を持つ	hasPart	あれば必須	0以上	
50	参照される	isReferencedBy	あれば必須	0以上	
51	参照する	references	あれば必須	0以上	
52	別フォーマットである	isFormatOf	あれば必須	0以上	
53	別フォーマットあり	hasFormat	あれば必須	0以上	
55	空間的	spatial			
56	国立情報学研究所 メタデータ主題語彙集 (地域)	NIIspatial		0~1	・JaLCシステムには、次の優先順位で初出のみ取り込まれる。 spatial > NIIspatial
59	権利	rights		0以上	

DataCite DOIを付与できるコンテンツ (NIItype)

データ・データベース、ソフトウェア

必須項目

- ・ タイトル (lang属性必須)
- ・ 作成者 (lang属性 = en推奨)
- ・ 公開者 (lang属性 = en推奨)
- ・ NII Type
- ・ URI
- ・ 本文URL
- ・ JaLC DOI (ra属性 = DataCite)

# Junii2 : 実際のデータ

```
<junii2 xmlns="http://irdb.nii.ac.jp/oai" xsi:schemaLocation="http://irdb.nii.ac.jp/oai http://irdb.nii.ac.jp/oai/junii2-3-1.xsd">  
<title lang="en">Naturalized plants in Japan : Seed-Image database</title>  
<alternative>日本の雑草 種子画像データベース</alternative>  
<creator lang="en">Yamashita, Jun</creator>  
<description>Naturalized Plants in Japan Seed-Image database is an image database of naturalized alien plant seeds recorded in Japan. This database contains plant images, information on morphological features of seeds and the name list of naturalized alien plants, as well as seed images. Seeds are categorized into 5 groups based on morphological features such as reticulation (wrinkles or projections), ridges, edges and achene fruit. Users can search seed images by family names, species names and those features.</description>  
<description>日本の雑草の種子画像を公開しています。帰化植物種子画像データベースと日本の雑草種子画像データベースを合併しました。</description>  
<publisher lang="en">Group of Wild Plant Science, Institute of Plant Science and Resources, Okayama University</publisher>  
<NIIttype>Data or Dataset</NIIttype>  
<URI>https://ousar.lib.okayama-u.ac.jp/63011</URI>  
<fullTextURL>https://ousar.lib.okayama-u.ac.jp/files/public/6/63011/20211220091442999252/Seed_DB.xlsx</fullTextURL>  
<selfDOI ra="DataCite">info:doi/10.48696/63011</selfDOI>  
<dateofissued>2021-12-20</dateofissued>  
<language>eng</language>  
</junii2>
```

# メタデータの課題（岡山大学）

- システム
  - メタデータ項目にlang属性がない  
タイトル、作成者、公開者 → 少し強引にマッピング（languageで判断）
- Junii2
  - タイトルを繰り返せない  
→ 複数言語（日・英）をdataciteに提供できない
  - 著者の所属、各種識別子等を提供できない

# メタデータ : JPCOARスキーマ

識別子	項目名	JPCOAR 要素名	属性	必須	繰り返し可否	形式等
1	タイトル	dc:title	*	必須	1以上	※lang属性は必須。
2.1	作成者識別子	jcoar:nameIdentifier	nameIdentifierScheme	あれば必須	0以上	同一著書でORCIDが複数ある場合、JALOSシステムには、初出のみが取り込まれる。
3.2	作成者姓	jcoar:familyName	*		0以上	※複数ある場合、JALOSシステムには、次の優先順位で1つだけ取り込まれる。 en<lang>を以てJALOSシステムへの登録が最も古いもの ・ない場合、JALOSシステムには、次の優先順位で取り込まれる。 jcoar:creatorNameの「姓」部分>jcoar:creatorName
3.4	作成者名	jcoar:givenName	*	必須	1以上	※複数ある場合、JALOSシステムには、次の優先順位で1つだけ取り込まれる。 en<lang>を以てJALOSシステムへの登録が最も古いもの ・ない場合、JALOSシステムには、次の優先順位で取り込まれる。 jcoar:creatorNameの「名」部分>jcoar:creatorName>jcoar:familyName
3.2	所属機関名	jcoar:affiliationName	*	あれば必須	0以上	※複数ある場合、JALOSシステムには、次の優先順位で1つだけ取り込まれる。 en<lang>を以てJALOSシステムへの登録が最も古いもの
4	著者	jcoar:contributor	contributorType	あれば必須	0以上	※lang属性、contributorTypeは必須。
4.1	著者識別子	jcoar:nameIdentifier	nameIdentifierScheme	あれば必須	0以上	同一著書でORCIDが複数ある場合、JALOSシステムには、初出のみが取り込まれる。
4.2	著者姓	jcoar:familyName	*		0以上	※複数ある場合、JALOSシステムには、次の優先順位で1つだけ取り込まれる。 en<lang>を以てJALOSシステムへの登録が最も古いもの ・ない場合、JALOSシステムには、次の優先順位で取り込まれる。 jcoar:contributorNameの「姓」部分>jcoar:contributorName
4.4	著者名	jcoar:givenName	*	あれば必須	0以上	※複数ある場合、JALOSシステムには、次の優先順位で1つだけ取り込まれる。 en<lang>を以てJALOSシステムへの登録が最も古いもの ・ない場合、JALOSシステムには、次の優先順位で取り込まれる。 jcoar:contributorNameの「名」部分>jcoar:contributorName>jcoar:familyName
4.2	所属機関名	jcoar:affiliationName	*	あれば必須	0以上	※複数ある場合、JALOSシステムには、次の優先順位で1つだけ取り込まれる。 en<lang>を以てJALOSシステムへの登録が最も古いもの
7	権利情報	dc:rights	rightsScheme	あれば必須	0以上	
9	分類	jcoar:subject	subjectScheme	あれば必須	0以上	
10	内容記述	datacite:description	descriptionType		0以上	
11	出版者	dc:publisher	en (推奨)	必須	1	・複数ある場合、JALOSシステムには、初出のみが取り込まれる。
12	日付	datacite:date	dataTypes="Issued"	あれば必須	0-1	・複数ある場合、JALOSシステムには、初出のみが取り込まれる。 ・JPCOARスキーマガイドラインに準拠しない場合及び空欄の場合は、JALOSシステムにて発行年に「9999」と設定される。
			dataType	あれば必須	0以上	・ある場合、dataTypeは必須。 ・dataTypes="Issued"以外、および2つ目以降のdataTypes="Issued"は、発行日ではなく、その他の日付として扱われる。
13	言語	dc:language			0-1	・複数ある場合、JALOSシステムには、初出のみが取り込まれる。 ・空欄の場合、または、対応する言語コードがJALOSシステムにない場合は、JALOSシステムにてurlと設定される。
14	資源タイプ	dc:type		必須	1	・dc:typeが以下のいずれかの場合、JALOSシステム上、「研究データ」としてメタデータ種別で取り込まれる。 ・dataset ・software
17	識別子	jcoar:identifier	identifierType="HDL" identifierType="URI"	必須	1	・複数ある場合、JALOSシステムには、identifierTypeを基に次の優先順位で取り込まれる。
18	ID登録	jcoar:identifierRegistration	identifierType="DataCite"	必須	1	・平角の"a-z","A-Z","0-9","_","/ "のみ使用可能。 ・プレフィックスは自動的に割り当てられたものに限り使用できる。
19	バージョン	datacite:revision		あれば必須	0-1	
19	関連情報	jcoar:relation	relationType	あれば必須	0以上	
19.1	関連識別子	jcoar:relatedIdentifier	identifierType="DOI" identifierType="ARK" identifierType="HDL" identifierType="PURL" identifierType="URI"	あれば必須	0以上	
21.1	位置情報(点)	datacite:geolocationPoint		※1ある場合、 いずれか必須	0-1	
21.1.1	経度	datacite:geolocationLongitude			0-1	
21.1.2	緯度	datacite:geolocationLatitude			0-1	
21.2	位置情報(空間)	datacite:geolocationBox		※1ある場合、 いずれか必須	0-1	
21.2.1	西部経度	datacite:westBoundLongitude			0-1	
21.2.2	東部経度	datacite:eastBoundLongitude			0-1	
21.2.3	南部緯度	datacite:southBoundLatitude			0-1	
21.2.4	北部緯度	datacite:northBoundLatitude			0-1	
21.3	位置情報(自由記述)	datacite:geolocationPlace		※1ある場合、 いずれか必須	0-1	
22.1	外部識別子	datacite:externalIdentifier	externalIdentifierType		0以上	
22.2	外部機関名	jcoar:externalName	en (推奨)	あれば必須	0以上	
23	研究種別	datacite:researchNumber			0以上	
25.1	本文URL	jcoar:url		必須	1以上	
25.2	フォーマット	jcoar:mimeType			0以上	
26.1	サイズ	jcoar:fileSize			0以上	

DataCite DOIを付与できるコンテンツ (dc:type)

データセット、ソフトウェア

必須項目

- ・ タイトル (lang属性必須)
- ・ 作成者名
- ・ 出版社 (lang属性 = en推奨)
- ・ 資源Type
- ・ 識別子 (identifierType=HDL or URI)
- ・ ID登録 (identifierType=DataCite)
- ・ 本文URL

Junii2との主な違い

タイトル複数言語OK、著者所属OK、各種識別子OK

# DataCiteに登録すると（JaLCとの違い）

- DataCite Search（<https://search.datacite.org/>）で検索可能  
→ DataCite Commonsに
- Google Dataset Search  
（<https://datasetsearch.research.google.com/>）で検索可能
- OpenAire（<https://www.openaire.eu/>）の収録対象
- Data Citation Indexの収録対象

# CiNii Research と DataCite Commons

The screenshot shows the CiNii Research interface. At the top, there is a navigation bar with tabs for 'すべて' (4), '研究データ' (1), '論文' (2), '本' (0), '博士論文' (0), and 'プロジェクト' (1). The '研究データ' tab is highlighted with a red box. Below the navigation bar, there is a red notification banner: '【10/4更新】2022年4月1日からのCiNii Articles CiNiiのサービスに関するアンケ'. The main content area displays the dataset entry for 'Naturalized plants in Japan : Seed-Image database' by Yamashita Jun. It includes a 'DOI' button, the author's name 'Yamashita Jun' (作成者), and a 'メタデータ' (Metadata) section with the following details: フォーマット: application/vnd.ms-excel, 公開日: 2021-12-20, 公開者: Group of Wild Plant Science, Institute of Plant Science and Resources, Okayama University. A '抄録' (Abstract) section at the bottom states: 'On this database, there are Seed-images of naturalized alien plants in Japan.'

<https://cir.nii.ac.jp/ja>

参考：DOI登録数（JaLC）：9,328,321

The screenshot shows the DataCite Commons interface. The URL bar displays '10.48696/63011'. The navigation bar includes 'Works', 'People', and 'Organizations'. The main content area displays the dataset entry for 'Naturalized plants in Japan : Seed-Image database' by Jun Yamashita. It includes the text 'Dataset published 2021 in Japan Link Center' and 'On this database, there are Seed-images of naturalized alien plants in Japan.' Under the 'Other Identifiers' section, it lists 'OAIPMH: oai:irdb.nii.ac.jp:01081:0005203073' and 'DOI registered December 25, 2021 via DataCite.' There are buttons for 'Dataset' and 'English', and a DOI link: <https://doi.org/10.48696/63011>.

<https://commons.datacite.org/>

DOI登録数（DataCite）299,383,196  
内JaLC 2,747

# FAIR原則

- F1. (メタ) データが、**グローバルに一意で永続的な識別子 (ID)** を有すること。
- F2. データがメタデータによって**十分に記述**されていること。
- F3. (メタ) データが検索可能なリソースとして、登録もしくは**インデックス化**されていること。
- F4. メタデータが、データの識別子 (ID) を明記していること。

# DataCite Metadata Schema v.4.3

Table 1: DataCite Mandatory Properties

<i>ID</i>	<i>Property</i>	<i>Obligation</i>
1	Identifier (with mandatory type sub-property)	M
2	Creator (with optional given name, family name, name identifier and affiliation sub-properties)	M
3	Title (with optional type sub-properties)	M
4	Publisher	M
5	PublicationYear	M
10	ResourceType (with mandatory general type description sub-property)	M

メタデータスキーマは、

**作成者、タイトル、発行者、発行年、リソースタイプ、識別子の6つの必須プロパティで構成**され、データの引用をサポート。推奨あるいは任意のさまざまなプロパティが用意されており、オブジェクトを記述・同定するために使用可能。

**推奨プロパティのうち、最も重要なのはDescriptionプロパティ**で、推奨サブプロパティのdescriptionType = "Abstract"とともに使用

DataCite Metadata Working Group. (2019). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 4.3. DataCite e.V.

<https://doi.org/10.14454/7xq3-zf69>

# データ引用原則の共同宣言

- 4. 識別

データを引用する際は、機械が使用でき、グローバルに一意で、かつ学術コミュニティで広く用いられている永続的な識別方法を記述すべきである。

- 6. 永続性

(識別子・メタデータが指し示す) データ自身の保存期間に関わらず、データの識別子、データを説明するメタデータを継続して保存すべきである。

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles.

Martone M. (ed.) San Diego CA: FORCE11; 2014 <https://doi.org/10.25490/a97f-egykh>