

研究データ:メタデータの共通項目への JaLCの令和5年度対応について

Japan Link Center (JaLC)

平尾義郎

2024年1月25日



1. 研究データ:メタデータの共通項目について
2. 新規追加予定の要素
3. 既存要素での運用
4. 例示の充実(マニュアル記述)
5. 出力ツール(REST API・OAI-PMH)の対応

1. 研究データ:メタデータの共通項目について
2. 新規追加予定の要素
3. 既存要素での運用
4. 例示の充実(マニュアル記述)
5. 出力ツール(REST API・OAI-PMH)の対応

研究データ:メタデータの共通項目 関連文書

内閣府「研究DX」サイト
<https://www8.cao.go.jp/cstp/kenkyudx>
 公的資金による研究データの管理・利活用

「公的資金による研究データの管理・利活用に関する基本的な考え方」
 (2021年4月 統合イノベーション戦略推進会議)
<https://www8.cao.go.jp/cstp/tyousakai/kokusaioopen/sank01.pdf>
 「我が国の公的資金による研究開発で共通的に用いられるべきと考えられるメタデータ共通項目を別紙のとおり定める。」別紙メタデータの共通項目

↓

メタデータの共通項目 最新版 2023年3月
https://www8.cao.go.jp/cstp/common_metadata_elements.pdf

ムーンショット型研究開発制度におけるメタデータ説明書 (第3.0版)
 2023年3月
https://www8.cao.go.jp/cstp/ms_metadatainstructions.pdf

メタデータ流通ガイドライン：研究データ編
 2023年3月3公開
 2023年11月更新
<https://ndlsearch.ndl.go.jp/guideline/researchdata>

メタデータ流通ガイドライン：研究データ編

メタデータ流通ガイドライン「研究データ編」(Metadata Distribution Guidelines「研究データ」)のメタデータにおいて採用される項目を説明します。どの資料種別でも共通的に採用される項目は、「共通編」で説明されており、「研究」は、「共通編」と重複しています。

10. 概略データ量

項目名 (日本語)	概略データ量
項目名 (英語)	File size
必須/任意	任意(選択式)
記述言語	—
備考	<ul style="list-style-type: none"> 管理対象データの概ねのデータ容量を以下から選択。 <ul style="list-style-type: none"> > <1GB > 1-10GB > 10-100GB > >100GB システムからデータ容量の値を出力できる場合は、データ容量の値そのものをセットしてもよい(GakuNin RDMまたはJAIRO Cloudを使用する場合は、データ容量の値の入力補完機能を提供)

目次

「研究データ編」について

- 項目説明
1. 助成情報
2. データNo
3. データの名称
4. 掲載日・掲載更新日
5. データの説明
6. データの分野
7. データ種別、データ量
8. データの提供方針
9. リポトリ情報
10. データ作成者
11. データ管理機関・管理者
12. 備考
13. メタデータのライセンス

「公的資金による研究データの管理・利活用に関する基本的な考え方」におけるメタデータの共通項目 (案)
 (2023年3月31日時点)

項目	必須/任意	備考
1 資金配分機関情報	必須※	公募型の研究資金を配分した資金配分機関(併呑含む)の英語略称 公募型の研究資金以外の場合は記入不要
2 体系的番号におけるプログラム情報コード	任意	公募型の研究資金の場合は、体系的番号のうち、「機関コード」および「施策・事業の特定コード」を表すコード 公募型の研究資金以外の場合は記入不要
3 プログラム名	任意	競争的研究費制度の名称
体系的番号	必須※	公募型の研究資金の場合は、研究費ごとに付与される体系的な番号
プロジェクト名	必須※	公募型の研究資金以外の場合は記入不要 プロジェクトの研究代表者が統括する研究開発の範囲の名称(e-Rad課題名称等) 公式名称がない場合は、研究者の所属機関のルールに従って入力
4 データNo.	必須	管理対象データを一意に特定するための番号 公募型の研究資金を配分した資金配分機関が付与 公募型の研究資金以外の場合はデータ管理機関(項目14)が付与
5 データの名称	必須	学会資料、報告資料、測定結果などの中身の分からない名称は避ける
6 掲載日・掲載更新日	必須	メタデータの掲載日・掲載更新日
7 データの説明	必須	掲載の中身の分かる内容を記載
8 データの分野	必須※	e-Radの研究分野(区分)を記載。e-Radとの連携により、自動入力される予定。
9 データ種別	必須	研究データ基盤システム上では、通常は「データセット」を標準とするが、データの特性に応じて「データセット」以外の種別を選択可能
10 概略データ量	任意	1GB未満、1GB以上100GB未満、100GB以上1000GB未満、1000GB以上の区分により記載
11 データの管理機関	必須	無償/有償、ライセンス情報、その他条件(引用の仕方等)を記載
12 データの公開予定日	必須	公開/共有/非公開/公開期間満了から選択
13 リポトリ情報	必須	公開期間満了を選択した場合、公開予定日を記載
14 データ管理者のe-Rad研究者番号	必須	管理者が所属する機関のe-Rad研究者番号
15 備考	任意	管理対象データを生み出した研究者の名前 管理対象データ作成者のe-Radの研究者番号

※公募型の研究資金による研究活動の場合

研究データ:メタデータの共通項目に対応する拡充を予定

「公的資金による研究データの管理・利活用に関する基本的な考え方」におけるメタデータの共通項目（案）
(2023年3月31日時点)

新規追加予定の要素

- 2. 体系的番号におけるプログラム情報コード、プログラム名
- 3. プロジェクト名
- 11. アクセス権
- 12. リポジトリ情報

既存要素での運用

- 14. データ管理者の連絡先

例示の充実(マニュアル記述)

- 8. データの分野
e-Rad研究分野の例示
- 9. データ種別
<resource_type type="Dataset">Dataset</resource_type>
の修正および
他のtype(サウンド、地図など)の例示追加
- 10. 概略データ量
<1GBおよび>100GBの例示

項目	必須/任意	備考
1 資金配分機関情報	必須※	公募型の研究資金を配分した資金配分機関（府省含む）の英語略称 公募型の研究資金以外の場合は記入不要
2 体系的番号におけるプログラム情報コード	任意	公募型の研究資金の場合は、体系的番号のうち、「機関コード」および「施策・事業の特 コード」を表すコード 公募型の研究資金以外の場合は記入不要
プログラム名	任意	競争的研究費制度の名称
体系的番号	必須※	公募型の研究資金の場合は、研究費ごとに付与される体系的な番号 公募型の研究資金以外の場合は記入不要
プロジェクト名	必須※	プロジェクトの研究代表者が統括する研究開発の範囲の名称（e-Rad課題名称等） 公式な名称がない場合は、研究者の所属機関のルールに従って入力
4 データNo.	必須	管理対象データを一意に特定するための番号 公募型の研究資金を配分した資金配分機関が付与 公募型の研究資金以外の場合はデータ管理機関（項目14）が付与
5 データの名称	必須	学会資料、報告資料、測定結果などの中身の分からない名称は避ける
6 掲載日・掲載更新日	必須	メタデータの掲載日・掲載更新日
7 データの説明	必須	端的かつ中身の分かる内容を記載
8 データの分野	必須※	e-Radの研究分野（主分野）。e-Radとの連携により、自動入力される予定。
9 データ種別	必須	研究データ基盤システム上では、通常は「データセット」を標準とするが、データの特性に 応じて「データセット」以外の種別を選択可能
10 概略データ量	任意	1GB未満、1GB以上10GB未満、10GB以上100GB未満、100GB以上等の区分により記載
管理対象データの利活用・提供方針	必須	無償/有償、ライセンス情報、その他条件（引用の仕方等）等を記載
11 アクセス権	必須	公開/共有/非共有・非公開/公開期間猶予から選択
公開予定日	必須	公開期間猶予を選択した場合、公開予定日を記載
リポジトリ情報	必須	現在のリポジトリ情報、あるいはプロジェクト後のリポジトリ情報
リポジトリURL・DOIリンク	任意	情報があれば記載
13 データ作成者	任意	管理対象データを生み出した研究者の名前
データ作成者のe-Rad研究者番号	任意	管理対象データ作成者のe-Radの研究者番号
データ管理機関	必須	各データを管理する研究開発を行う機関の法人名
データ管理機関コード	任意	データ管理機関のコード
データ管理者	必須	データ管理組織において各管理対象データを管理する担当者名
14 データ管理者のe-Rad研究者番号	任意	管理者のe-Radの研究者番号 e-Rad研究者番号がない管理者は記入不要、ある場合は番号研究者が番号を非公開にしたい場 合を除き必須
データ管理者の連絡先	必須	データ管理者の所属機関の住所や電話番号、メールアドレス等
15 備考	任意	

※公募型の研究資金による研究活動の場合

1. 研究データ:メタデータの共通項目について
2. 新規追加予定の要素
3. 既存要素での運用
4. 例示の充実(マニュアル記述)
5. 出力ツール(REST API・OAI-PMH)の対応

2.体系的番号におけるプログラム情報コード・プログラム名(1/2)

「メタデータ流通ガイドライン：研究データ編」からの抜粋

1-2. プログラム情報

メタデータ共通項目:「2.体系的番号におけるプログラム情報コード」

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
プログラム情報	体系的番号におけるプログラム情報コード	任意	jpcoar:fundingStreamIdentifier	fund_list/fund/award_number_group/program_id	dcterms:description
	プログラム名	任意	jpcoar:fundingStream	fund_list/fund/award_number_group/program_name	

- 体系的番号における「機関コード」および「施策・事業の特定コード」を合わせた英数字最大5桁のコード（体系的番号の構成は競争的研究費制度によって異なるため、該当するコードがない場合は空白）
- 公募型の研究資金以外の場合は記入不要

メタデータ共通項目:「2.プログラム情報名」

基本事項

- 各資金配分機関内の助成カテゴリ（助成制度・事業等）の名称を記入します。公募型の研究資金以外の場合は記入不要です。
- 制度情報の識別子には、「Crossref Funder」もしくは「体系的番号におけるプログラム情報コード」を選択します。

- 公募型の研究資金の場合は、体系的番号一覧に記載された競争的研究費制度の名称。公式な名称がない場合は、研究者の所属機関のルールに従って入力

体系的番号におけるプログラム情報コード

「体系的番号」の「機関コード」と「施策・事業の特定コード」を合わせたもの（例・科学技術振興機構（JST）のムーンショット型研究開発事業は、機関コードMJ、施策・事業の特定コードMSを合わせた「MJMS」を入力する）

2.体系的番号におけるプログラム情報コード・プログラム名(2/2)

➤ データ例(予定) :

【program_id】

【program_name】

制度情報の識別子としての「Crossref Funder」の例(JST ERATO)

```
<fund>
  <funder_name lang="en">Japan Science and Technology
    Agency</funder_name>
  <funder_identifier type="ROR">
    https://ror.org/00097mb19</funder_identifier>
  <funder_identifier type="FundRef">
    https://doi.org/10.13039/501100002241</funder_identifier>
  <award_number_group>
    <award_number>JPMJER1402</award_number>
    <project_name>齊藤スピ>量子整流プロジェクト</project_name>
    <program_id>https://doi.org/10.13039/501100009024
      </program_id>
    <program_name>ERATO</program_name>
  </award_number_group>
</fund>
```

制度情報の識別子としての「体系的番号におけるプログラム情報コード」の例(JST CREST)

```
<fund>
  <funder_name lang="en">Japan Science and Technology
    Agency</funder_name>
  <funder_identifier type="ROR">
    https://ror.org/00097mb19</funder_identifier>
  <funder_identifier type="FundRef">
    https://doi.org/10.13039/501100002241</funder_identifier>
  <award_number_group>
    <award_number>JPMJCR1604</award_number>
    <project_name>データ科学に基づく作物設計基盤技術の構築
      </project_name>
    <program_id>MJCR</program_id>
    <program_name>CREST</program_name>
  </award_number_group>
</fund>
```

➤適用状況 :

令和5年度中にシステム改修し、対応**予定**

3. プロジェクト名

「メタデータ流通ガイドライン：研究データ編」からの抜粋

1-3. 課題情報

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
課題情報(課題番号・課題名)	体系的番号	推奨	jpcoar:awardNumber	fund_list/fund/award_number_group/award_number	dcterms:description
	プロジェクト名	推奨	jpcoar:awardTitle	fund_list/fund/award_number_group/project_name	

基本事項

- 研究課題番号と、研究課題名を記載します。

メタデータ共通項目:「3.体系的番号」

- 公募型の研究資金の場合は、研究費ごとに付与される体系的番号。公募型の研究資金以外の場合は記入不要
- 「メタデータ共通項目」上は、公募型の研究資金による研究活動の場合必須項目となっています。

メタデータ共通項目:「3.プロジェクト名」

- プロジェクトの研究代表者が統括する研究開発の範囲の名称。e-Radの課題名称があればそれを入力します。
- 「メタデータ共通項目」上は、公募型の研究資金による研究活動の場合必須項目となっています。

➤データ例(予定):

【award_number】
【project_name】

```
<fund>
  <funder_name lang="en">Japan Science and Technology Agency</funder_name>
  <funder_identifier type="ROR">
    https://ror.org/00097mb19</funder_identifier>
  <funder_identifier type="FundRef">
    https://doi.org/10.13039/501100002241</funder_identifier>
  <award_number_group>
    <award_number>JPMJCR1604</award_number>
    <project_name>データ科学に基づく作物設計基盤技術の構築</project_name>
  </award_number_group>
</fund>
```

➤適用状況:

体系的番号は**対応済**、プロジェクト名については令和5年度中にシステム改修し**対応予定**

11. アクセス権

「メタデータ流通ガイドライン：研究データ編」からの抜粋

8-2. アクセス権

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
アクセス権	アクセス権	推奨	dcterms:accessRights	access_rights	dcterms:accessRights

基本事項

- 管理対象データのアクセス権について記録します。

メタデータ共通項目「11.アクセス権」

- 管理対象データのアクセス権について下記から選択
 - 公開 (open access)：一般に公開されるデータ
 - 共有 (restricted access)：限定された関係者での利活用を目的として共有されるデータ
 - 非共有・非公開 (metadata only access)：公開も共有もされないデータ
 - 公開期間猶予 (embargoed access)：非共有・非公開のデータのうち、一定期間後に公開されるデータ これを選択した場合は「公開予定日」の記載が必須
- 「メタデータ共通項目」上は必須項目となっています。

➤データ例(予定)：

【access_rights】

```
<access_rights date="20230402">
  embargoed access</access_rights>
```

embargoed accessの場合のみdate属性が必要

➤適用状況：

令和5年度中にシステム改修し、対応**予定**

12. リポジトリ情報

「メタデータ流通ガイドライン：研究データ編」からの抜粋

9-1. リポジトリ情報

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
リポジトリ名	リポジトリ情報	推奨	jpcoar:catalog /dc:title (irdbProvenance)	repository/repository_name_list /repository_name	/dcndl:BibAdminResource/dcndl:bib RecordSubCategory

(共通編 3. 出版・提供と一部重複)

基本事項

- メタデータを提供（作成）した機関またはデータベース等の情報を記録します。(※共通編「3-2メタデータ提供者」参照)

メタデータ共通項目:「12リポジトリ情報」

- 管理対象データの現在のリポジトリ情報、あるいはプロジェクト終了後のリポジトリ情報について記録します。
- 「メタデータ共通項目」上は必須項目となっています。

▶データ例(予定) :

【repository_name】

```

<repository>
  <repository_name_list>
    <repository_name lang="ja">東京大学学術資産等
    アーカイブズポータル
    </repository_name>
    <repository_name lang="en">UTokyo Academic
    Archives Portal
    </repository_name>
  </repository_name_list>
</repository>

```

▶適用状況 :

令和5年度中にシステム改修し、対応**予定**

1. 研究データ:メタデータの共通項目について
2. 新規追加予定の要素
3. 既存要素での運用
4. 例示の充実(マニュアル記述)
5. 出力ツール(REST API・OAI-PMH)の対応

14. データ管理者の連絡先

「メタデータ流通ガイドライン：研究データ編」からの抜粋

11-5. データ管理者の連絡先

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
データ管理者の連絡先	データ管理者の連絡先	推奨	jpcoar:contributor@contributorType="ContactPerson"	contributor_list/contributor@contributor_type="ContactPerson"/names/first_name	-

基本事項

- 当該研究データの管理者の連絡先を記録する項目です。

メタデータ共通項目: 「14. データ管理者の連絡先」

- 管理者の所属組織名、電話番号、メールアドレス等を記入（メールアドレスを必須、他を任意とする。メールアドレスがない場合は、所属組織名（法人内の部署名等）、住所、電話番号を必須とする）
- 当該研究データが非公開になっている場合に管理対象データに興味を有する第三者が必ずデータ管理者にたどり着けるようにする。
- 「メタデータ共通項目」上は必須項目となっています。

データ例(案) :

【first_name】

```
<contributor sequence="1" type="person" contributor_type="ContactPerson">
  <names lang="ja">
    <first_name> ジャパンリンクセンター info@japanlinkcenter.org
  </first_name>
  </names>
```

適用状況(予定) :

令和5年度中に研究データ情報登録マニュアルを改訂し次の内容を明記する予定。

「公的資金による研究データの管理・利活用に関する基本的な考え方」での研究データ、すなわち「公的資金による研究開発の過程で生み出される全てのデータ」の場合は、
 contributor_list/contributor@contributor_type="ContactPerson"/names/first_name
 にデータ管理者の所属組織名、所属組織の電話番号・メールアドレス等(業務アドレスやグループアドレス)を記入する。公開データとなるため、個人を特定する電話番号・メールアドレスは記入しない。

1. 研究データ:メタデータの共通項目について
2. 新規追加予定の要素
3. 既存要素での運用
4. 例示の充実(マニュアル記述)
5. 出力ツール(REST API・OAI-PMH)の対応

8. データの分野

「メタデータ流通ガイドライン：研究データ編」からの抜粋

6. データの分野

項番	項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
6	データの分野	データの分野	推奨	jpcoar:subject	subject_list/subject	dcterms:subject/rdf:Description@rdf:about dcterms:subject/rdf:Description/rdf:value

6-0. 概要

基本事項

- 研究プロジェクトもしくはデータの研究分野を記録します。
- 分野に関する情報が記録されていることで、利用者が情報資源の中身を推測することができ、また、発見可能性を高めることができます。

メタデータ共通項目：「8データの分野」

- e-Radで登録したプロジェクトの研究分野（主）に基づき、研究内容のいずれかを入力。データの分野とプロジェクトの研究分野が異なる場合は、データの分野を入力。
- e-Radの研究分野はe-Radのウェブサイト(CSVファイル) [\[2\]](#)を参照のこと

▶データ例：

【subject_list】

```
<subject_list>
```

```
<subject lang="ja" subject_scheme="NDC" scheme_uri="https://www.ndl.go.jp/jp/data/NDC10code202206.pdf">420物理学</subject>
```

```
<subject lang="ja">自然言語処理</subject>
```

```
<subject lang="ja">機械学習</subject>
```

```
<subject lang="ja" subject_scheme="e-Rad" scheme_uri="https://www.e-rad.go.jp/datasets/files/bunya_bunya.csv">ライフサイエンス</subject>
```

```
<subject lang="en" subject_scheme="e-Rad" scheme_uri="https://www.e-rad.go.jp/datasets/files/bunya_bunya.csv">Life Science</subject>
```

```
</subject_list>
```

※詳細は研究データ情報登録マニュアルのデポジットXMLメタデータ項目リスト参照

subject_list/subject

9. データの種別

「メタデータ流通ガイドライン：研究データ編」からの抜粋

7-1. データ種別

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
データ種別	データ種別	強く推奨	dc:type	resource_type	dcndl:materialType@rdf:resource

基本事項

- 研究データの種別に関する情報を記録します。採用するスキーマにより資料種別の情報を選択してください。

メタデータ共通項目:「9 データ種別」

- JPCOARスキーマの[資源タイプ語彙別表参照【Ver2.0】](#)から選択

➤データ例：

【resource_type】

データセットの例

```
<resource_type type="Dataset">調査データ</resource_type>
```

静止画の例

```
<resource_type type="Image">静止画</resource_type>
```

音声の例

```
<resource_type type="Sound">音声</resource_type>
```

ソフトウェアの例

```
<resource_type type="Software">ソースコード</resource_type>
```

地図の例

```
<resource_type type="Other">地図</resource_type>
```

※詳細は研究データ情報登録マニュアルのデポジットXMLメタデータ項目リスト参照

resource_type

10. 概略データ量

「メタデータ流通ガイドライン：研究データ編」からの抜粋

7-2. データ量

項目名	メタデータ共通項目	推奨度	JPCOARスキーマ	DC_JaLC	DC-NDL (RDF)
データ量	概略データ量	任意	jpcoar:extent	size_list/size	dcterms:extent

基本事項

- 当該情報資源のデータサイズを記録する項目です。

メタデータ共通項目:「10概略データ量」

- 管理対象データの概ねのデータ容量を以下の候補から選択
 - <1GB
 - 1-10GB
 - 10-100GB
 - >100GB
- システムからデータ容量の値を出力できる場合は、データ容量の値そのものをセットしてもよい。

➤データ例：

【size】

240MBの例

```
<size_list>
<size>240MB</size>
</size_list>
```

【<1GB】の例

```
<size_list>
<size>&lt;1GB </size>
</size_list>
```

1-10GBの例

```
<size_list>
<size>1-10GB </size>
</size_list>
```

【>100GB】の例

```
<size_list>
<size>&gt;100GB</size>
</size_list>
```

※詳細は研究データ情報登録マニュアルのデポジットXMLメタデータ項目リスト参照

size_list/size

「<」や「>」はそのまま入力できず、それぞれ「<」と「>」という文字列で入力する必要がある(XMLアップロードの場合)。

1. 研究データ:メタデータの共通項目について
2. 新規追加予定の要素
3. 既存要素での運用
4. 例示の充実(マニュアル記述)
5. 出力ツール(REST API・OAI-PMH)の対応

新規要素(プログラムID、プログラム名、プロジェクト名、アクセス権、リポジトリ情報)については、以下のようにv2を利用することで取得可能となる。v2を指定しない場合は従来通りの出力。

- DOIのメタデータ取得
例) <https://api.japanlinkcenter.org/v2/10.11502/example-doi>

OAI-PMHのmetadataPrefixパラメータ値に「DC_JaLC_v2」を新規追加。リクエスト時のmetadataPrefixパラメータに「DC_JaLC_v2」を指定することで上述の新規要素が取得可能となる。従来通りの「DC_JaLC」を指定すると従来通りの出力。

- GetRecord によるデータ取得例
 - https://japanlinkcenter.org/jalc/infoProvide/oaiPmh/index?verb=GetRecord&metadataPrefix=DC_JaLC_v2&set=RD&identifier=10.11502/example-doi
 - 注: 改修リリース後、半月程度後に「DC_JaLC_v2」が利用可能になる予定。

ありがとうございます