

機関リポジトリへの 研究データ登録ガイドライン

Ver. 1.0



目次

I.	本ガイドラインについて	3
II.	事前知識	5
1.	登録の前に.....	5
2.	「研究データ」として登録対象となる資源.....	5
2-1.	研究データの定義	5
2-2.	論文と根拠データ	5
III.	リポジトリ登録.....	7
1.	リポジトリ登録のパターン	7
パターン 1	研究データのみを単体で登録する.....	7
パターン 2	研究データを論文の一部として登録する場合	10
2.	登録依頼.....	10
2-1.	登録データの受領方式	10
2-2.	提出物の確認作業	12
3.	登録作業.....	13
コラム「研究データ」登録の際に登録必須となる項目		15
4.	登録完了後	18
4-1.	確認作業.....	18
4-2.	依頼者への完了通知	18
コラム「データリポジトリにおける処理」.....		19
データに対する処理.....		19
データの再構築		20

IV.	研究データの内容・メタデータ更新.....	22
1.	更新準備	22
2.	更新作業	22
	ファイルの更新.....	22
3.	更新完了後.....	23
V.	研究データの削除・非公開	24
1.	研究データの非公開.....	24
1-1.	概要	24
1-2.	対応方法	24
2.	研究データの削除.....	24
2-1.	概要	24
2-2.	対応方法.....	25
VI.	登録後の対応	27
1.	統計情報の取得.....	27
	登録パターン 1 の場合.....	27
	登録パターン 2 の場合.....	27
2.	問い合わせ内容の記録管理.....	27

改訂履歴

版数	発行日	改訂履歴
Ver. 1.0	2025 年 9 月 5 日	初版公開

I. 本ガイドラインについて

本ガイドラインは、オープンアクセスリポジトリ推進協会(JPCOAR)研究データ作業部会によって作成された、機関リポジトリへ研究データを登録する際に有用な情報をまとめたガイドラインである。国内の大学・研究機関内で生産された学術情報の積極的な発信を担う基盤としてスタートした機関リポジトリ^[1]は、2005年の最先端学術情報基盤(CSI)委託事業(国立情報学研究所)^[2]や、2012年のJAIRO Cloud サービス開始^[3]などを経て、学術情報流通における一定のポジションを確立しつつある。より最近では、2024年2月に公開された「学術論文等の即時オープンアクセスの実現に向けた基本方針」において、公的資金による学術雑誌論文及びその根拠データを即時に機関リポジトリ等へ掲載することが義務付けられた^[4]。こういった政策的な動きを受け、各大学・研究機関は自機関における学術情報の発信戦略において、機関リポジトリの活用を前提に業務を組み立てる必要に迫られている。

一方で、機関リポジトリの活用は各大学・研究機関によって異なるモチベーションで進められてきており、政策的な要求とコンテンツの収集実態には大きな乖離がある^[5]。中でも、学術雑誌論文の根拠データについては標準的な登録ルールが確立されておらず^[6]、実務的に参照できる文献が管見の限り国内に見当たらない。そこで本ガイドラインでは、機関リポジトリへの研究データ登録手順を一連のフローとして整理するとともに、各ステップで留意すべきポイントを国内機関における事例を参考にしてまとめている。各大学・研究機関の機関リポジトリ担当者は、本ガイドラインで取り上げられている項目を参照することで、研究データ登録における業務フローの全体像を手続きのレベルで理解することが可能になる。さらに、具体的な業務マニュアル作成・改訂にあたっては、本ガイドラインを自機関の事情に合わせてカスタマイズしながら利用することも可能である。

本ガイドラインは、2023年度に実施した国内複数機関へのヒアリングをもとに全体の構成を整理したうえで、各研究分野におけるデータキュレーション活動を整理・分析^[7]し、国内における機関リポジトリ業務向けに再構築したものである。また、メタデータ登録に当たっては、内閣府 科学技術・イノベーション推進事務局が公表している「公的資金による研究データの管理・利活用に関するメタデータ説明書 第1.0版」^[8]に対応した内容となっている。

^[1] [科学技術・学術審議会研究計画・評価分科会情報科学技術委員会デジタル研究情報基盤ワーキング・グループ]:学術情報の流通基盤の充実について(審議のまとめ), https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu2/toushin/image/020401c.pdf

^[2] 国立情報学研究所:学術機関リポジトリ構築連携支援事業.
<https://www.nii.ac.jp/irp/rfp/>

^[3] オープンアクセスリポジトリ推進協会:JAIRO Cloud.
<https://jpcoar.repo.nii.ac.jp/page/42>

- [4] 統合イノベーション戦略推進会議:学術論文等の即時オープンアクセスの実現に向けた基本方針. https://www8.cao.go.jp/cstp/oa_240216.pdf
- [5] 田辺浩介, 有馬良一, 岩井雅史, 佐藤知生, 三上絢子, 結城憲司, 南山泰之:機関リポジトリへの論文及び根拠データ掲載に関する実態調査報告. 情報知識学会誌, Vol.34, No.2, pp. 127-134, 2024.
- [6] 三上 絢子, 南山 泰之:論文根拠データの公開状況測定に向けた調査手法の確立. 情報知識学会誌, Vol.33, No.2, pp. 242-247, 2023.
- [7] Minamiyama Y, Takeda H, Hayashi M, Asaoka M, Yamaji K (2024) A study on formalizing the knowledge of data curation activities across different fields. PLOS ONE 19(4): e0301772.
<https://doi.org/10.1371/journal.pone.0301772>
- [8] 内閣府 科学技術・イノベーション推進事務局. “公的資金による研究データの管理・利活用に関するメタデータ説明書 第 1.0 版”.
<https://www8.cao.go.jp/cstp/metadainstructions.pdf>

II. 事前知識

1. 登録の前に

研究データを登録する際には、本マニュアルのほかに、

- ・ 自機関の機関リポジトリの運用指針
(例) [京都大学学術情報リポジトリ運用指針](#)
- ・ 自機関の研究データ管理に関するポリシーやガイドライン
(例) [国内大学の研究データポリシー\(一覧\)](#)

等についても、適宜参照のこと。

機関リポジトリへの研究データ登録においては、予め大学・機関で定めている「研究データポリシー」をもとに、研究データの管理方法や公開及び共有の基準様々など、様々な判断を行う。

各大学・研究機関では、原則として 2025 年までに研究データポリシーを定めることとなっている。

2. 「研究データ」として登録対象となる資源

2-1. 研究データの定義

「研究データ」が指し示す範囲は研究分野ごとに異なる。本マニュアルでは、リポジトリへの登録を前提とするため、電磁的な方法により管理が可能なデータのうち、研究成果として公開するデータを対象とする。

「研究データ」には、数値、テキスト、画像、音声、動画など、さまざまな形式があり、データが利用される文脈により、論文の根拠データ、ソースデータ、派生データなど様々な呼称が用いられることがある。

(参考) “研究データの公開・利用条件表示ガイドライン ver.1.0”. 研究データ利活用協議会 (RDUF) 研究データライセンス小委員会, 2019, 32p. https://doi.org/10.11502/rdof_license_guideline, (参照 2025-09-03).

2-2. 論文と根拠データ

公的資金を受けて作成した研究データ(公開可能)について、機関リポジトリ等の登録が進んでいる。また、即時オープンアクセス政策により、論文の「根拠データ」についても対応が必要となった。

論文の「根拠データ」も「研究データ」に含まれるものであり、原則として「研究データ」の登録方法に従って登録する。(加えて、論文との関係を登録時に入力する)

なお、「根拠データ」とは「掲載電子ジャーナルの執筆要領、出版規程等において、透明性や再現性確保の観点から必要とされ、公表が求められる研究データ」と位置づけられている。

※参考資料

- ・ [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](#)
- ・ [学術論文等の即時オープンアクセスの実現に向けた基本方針](#)

III. リポジトリ登録

1. リポジトリ登録のパターン

研究データ登録においてはいくつかのパターンが考えられる。

- ・ パターン 1 : 研究データのみを単体で登録する
- ・ パターン 2 : 研究データを論文の一部として登録する。

いずれの方式を用いるかは、(出版社のデータポリシー、分野の慣行、添付する研究データの規模、位置づけ等)各々のケースによる。

パターン 1 研究データのみを単体で登録する

このパターンを選択する場面としては、以下のようなときである。

- ・ 論文発表から独立して研究データを作成し、成果として利活用に供したいとき
- ・ 関連論文を発表しており、かつ研究データも別途登録する場合(両アイテム間の関係は別途記述)

論文の根拠データをこのパターンで登録する場合は、「**関連情報**」のスキーマ項目を用いて関係を表現すること

(参考)[関連情報 | JPCOAR スキーマガイドライン](#)

参考例 -----

※以下のスクリーンショット画像はいずれも 2025 年 8 月時点のものです。

研究データが単体で登録された例

京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

京都大学
KYOTO UNIVERSITY

ホーム > 350 生存圏研究所 > 研究データ > コンピュータービジョン…

ダウンロード数: 27

このアイテムの引用には次の識別子を使用してください: <https://doi.org/10.14989/287077>

コンピュータービジョンによる木材細胞の変形解析に使用したデータセット

ファイル [Download All](#)

[dataset.zip \(885.2 MB\)](#) 記述: [Download](#)

MD5: b22e382946d265bfafc78bf1f2b270b5

発行日
2024-02-22

その他のタイトル
Dataset for analyzing cell deformation by computer vision

著者
陳, 碩也
KAKEN 研究者検索
杉山, 淳司
KAKEN 研究者検索

キーワード
細胞変形解析, コンピュータービジョン, cell deformation analysis, computer vision

目次
dataset.zip ファイルに以下の内容が含まれます。
1. フォルダー "large_image_sample": 木材試料の3点曲げ試験におけるその断面の連続画像 (0 sec .img ~ 203.sec .img)

<http://hdl.handle.net/2433/287077>

関係メタデータで当該データを根拠にした論文へのリンクを示す。

dc.relation.isSupplementTo

<https://doi.org/10.1515/hf-2022-0096>

リンク先の成果論文

京都大学学術情報リポジトリ
KURENAI
Kyoto University Research Information Repository


京都大学
KYOTO UNIVERSITY

ホーム > 100 農学研究科・農学部 > 学術雑誌掲載論文等 > Flexural behavior of woo...

ダウンロード数: 95

このアイテムの引用には次の識別子を使用してください: <http://hdl.handle.net/2433/285048>

Flexural behavior of wood in the transverse direction investigated using novel computer vision and machine learning approach

ファイル

Download All

hf-2022-0096.pdf (4.5 MB)

記述: Download

発行日

2022-10-26

キーワード

cell deformation, deep learning, flexural behavior, individual cell tracking, semantic segmentation

著者

Chen, Shuoye

KAKEN 研究者検索

ORCID iD

 0000-0001-6950-1856 (unconfirmed)

Awano, Tatsuya

KAKEN 研究者検索

ORCID iD

 0000-0002-4735-4506 (unconfirmed)

Yoshinaga, Arata

KAKEN 研究者検索

ORCID iD

 0000-0002-2935-3067 (unconfirmed)

Sugiyama, Junji

ORCID iD

 0000-0002-5388-4925 (unconfirmed)

誌名

Holzforschung

巻

76

号

10

開始ページ

875

終了ページ

885

抄録

<http://hdl.handle.net/2433/285048>

成果論文のメタデータも根拠データへのリンクを記載できる

dc.relation.isSupplementedBy

<https://doi.org/10.14989/287077>

(参考)

JPCOAR スキーマガイドラインでは、論文と研究データとを別アイテムとして登録した上で、関係を表すメタデータを記述する方式を推奨している。

パターン 2 研究データを論文の一部として登録する場合

このパターンを選択した場合、論文の付録物として登録するため、CiNii Research では論文扱いとなる。そのため、「研究データ」として CiNii Research 等で検索・カウントされなくなること
に留意する必要がある

詳細は CiNii Research の[「研究データ」「根拠データ」の収録について](#)を参照すること。

また、「研究データ」に関するメタデータは記述できない

2. 登録依頼

2-1. 登録データの受領方式

各大学・研究機関で定められた方式で申請を受理し、データを受領する。登録申請とデータの提出は、メールベースで行われる場合と専用のシステムで行われる場合がある。

A. メールベースで行われる場合

メールベースで行われる場合は、申請書を準備して提出内容を整理しているケース、メールに直接必要な情報を書きこんでもらうケースがある。データの受渡し方法は、メール添付、ストレージへのアップロードなどが考えられる。

B. 専用のシステムの場合

登録データの受領を専用のシステムで行う場合は、当該システムのマニュアルに従う。

以下、提出物別に注意点を列挙する。

研究データ本体

- ・ 論文の根拠データであるか独立したデータセットであるか
- ・ 対応可能なデータの上限(上限を超えるデータへの対応)
- ・ 長期的な保存に適したフォーマットか(推奨:csv、.txt、.xml、.tiff、.mp4、.pdf (PDF/A)、etc.)
- ・ 研究データを適切に確認、利用するに十分な情報を「readme.txt」ファイルとして含めているか。

説明文書

データ文書は、寄託データの内容や作成方法、利用方法を記述した文書である。以下の情報が含まれることが望ましい:

- ・ データの概要(データの目的、収集方法、作成者)
- ・ データの構成(ファイル形式、データ構造、変数の説明)

- ・ データの前処理や加工手順
- ・ 利用に関する注意事項(制約事項、ライセンス情報など)

コード

コードは、データの処理や分析を再現するものである。以下のような情報が含まれることが望ましい。

- ・ データ処理・解析のためのスクリプト(Python, R など)
- ・ 実行環境(使用ライブラリ、バージョン情報)
- ・ 再現手順(どの順序でコードを実行するかの説明)
- ・ 出力ファイルの説明(生成されるファイルや可視化結果など)

メタデータ

機関リポジトリで扱われるメタデータには様々な種類があるが、概ね以下のような内容が記述されていることが望ましい。

- ・ 基本情報
データタイトル、作成者、所属機関、作成年月日
- ・ データの内容説明
概要、キーワード、関連研究
- ・ データアクセス情報
公開・非公開設定、ライセンス、DOI など
- ・ 技術的情報
ファイル形式、データサイズ、保存方法
- ・ 関係性情報
関連論文、プロジェクト、他のデータセットとの関係

機関リポジトリへ一括でメタデータを取り込む際には、所定の一括登録フォーマットに沿ってメタデータを準備してもらう方式が簡便である。Excel でフォーマットを用意する場合は、セル結合を使用しない、入力制限を行うなど機械処理しやすいよう配慮する必要がある。

登録許諾書

登録許諾書は、データ寄託者(データ提供者)とリポジトリ運営者の間で、データの登録・公開・利用に関する権利と責任を明確にするための文書である。データ登録許諾書には、主に以下のような項目が含まれる。

- ・ データのオーナーシップ
寄託者がデータの保有者として適切な権利を有しており、リポジトリに登録する権限を持つことを確認する。
- ・ データの登録および保存

リポジトリがデータを長期的に保存・管理し、適切なアクセス権限のもとで提供することを許諾する。

- ・ ライセンスと利用条件
データがどのような条件で公開・利用されるかを明確にする。(例:CC BY、CC0、独自ライセンス)
- ・ データの変更・削除
寄託者がデータの修正・更新を希望する場合の手続きや、削除が可能かどうかについて説明する。
- ・ 責任の範囲
データの正確性や法的問題については寄託者が責任を負い、リポジトリ側はデータの適切な管理を行うが、利用者のデータ利用結果に関しては責任を負わないことを明記する。

2-2. 提出物の確認作業

提出物に不足がないか確認する。また、自機関がデータ登録にあたり案内している条件に合致しているか確認する。主なポイントは以下のとおり。

寄託者の認証・認可

寄託者の認証方法は、リポジトリ担当者がその場で確認する方法(例:大学ドメインのメールアドレスからデータを受領した場合)と、事前に登録用アカウントを作成してもらう場合がある。担当者間で寄託者の身元を確認する手順を共有しておくといよい。

登録ポリシー確認

- ・ 研究成果の根拠となるデータか
「根拠データ」の場合、対応する論文等のアイテムについて特定する。「論文」「根拠データ」双方のアイテムについて、関連情報(jpcoar:relation)を記載する。(relationTypeは”isSupplementedBy”, ”isSupplementTo”)
関係を記載して登録すると、リポジトリや CiNii Research で検索された場合に、成果の全体像が把握でき、発信力強化、成果アピールに繋がる。
- ・ 根拠データにとどまらず、研究データは積極的に登録・公開する。ただし、オープン・アンド・クローズ戦略に基づき、研究分野等の特性や、データを管理する組織の特性に配慮して、研究者自身が「公開」、「共有」又は「非共有・非公開」の判断を行う。
- ・ 公募型の研究資金を受けて作成されたものか否か
前者の場合、メタデータ共通項目の必須項目を登録する必要がある。依頼者から提供された情報に不足が無い確認すること。

テクニカルチェック

提出物の内容に不備がないか確認する。

登録許諾書のレビュー

データ公開にあたって研究者側の責任で確認いただく事項について触れておく必要がある。

- 公開に際してデータに権利を持つ全員の合意を得ているか、共同研究において機密契約などの事項はないか。
- 当該研究データに、個人情報や安全保障輸出管理等の法的問題にかかわる事項、出願予定の発明等、及び企業の秘密情報が含まれないか。
- 外部資金等による研究プロジェクトで、契約書等で研究データの取り扱いについて定めがある場合は、公開に制約や条件がないか。
- 当該分野の研究コミュニティでの法的および倫理的要件等により、一般公開に制約がある内容を含まないか。

※データの利活用にあたり制約事項がある場合、メタデータ共通項目「管理対象データの利活用・提供方針」にも当該記載があることを確認する

3. 登録作業

(1) 研究データのフォーマットタイプについて、以下のような事項を確認する。

- リポジトリに登録できるフォーマットの制限がないか確認する。
- 見慣れない拡張子のファイルを受領したら、以下の Web ページなどを参照して確認する。
[よくある MIME タイプ](#)
- 基本的には、申請者から提出されたファイルをそのまま搭載する。セキュリティレベルも変更しない。
- ファイルを開いて読み取りエラーが起こらないか確認する。特殊なソフトウェアが必要で、手元では開けない場合は、確認を省略してよい。
- データに不備がないか、利活用に支障がないか、「[データリポジトリにおける処理](#)」を参照して確認する。

(2) ファイルサイズを確認し、各機関で定めた上限あるいはリポジトリシステムの登録容量の上限を超えるようなファイルの分割等を検討する。

- 大きなデータは登録に時間がかかる。場合によってはそのままフリーズしてしまうことがあるので、適切なサイズに分割する。その場合のファイル名の命名 ルールを決めておくといよい。

- 各大学のポリシーによるが、本文ファイルは必ずしも機関リポジトリ上に置く必要は無い。外部リポジトリに置き、その URL を記載する方法も可能である。

(3) 登録すべき分類(インデックスやカテゴリー)を確認する。ない場合は作成する。

(4) メタデータおよび研究データを登録する。

【コラム「研究データ」登録の際に必須となる項目】も参照のこと。

- 登録希望者から提出された申請内容に準じてメタデータを入力し、ランディングページを作成する。
- 研究データ登録申請書がある場合は、その項目に応じてシステムにテンプレートを作成しておくといい。
- 申請内容と相違がないか、メタデータの最終チェックを行う。可能であれば、例えば readme にのみ記載がある情報がないかも確認し、補完しておくとい。
- 登録する研究データのファイルを選択して登録する。

(5) DOI の付与希望があれば、DOI を付与する。

- JaLC のエラーチェックや IRDB のエラーチェックに注意する。

参考:

■ <https://support.irdb.nii.ac.jp/ja/harvest/jpcoar/validation>

■ https://support.irdb.nii.ac.jp/ja/harvest/jpcoar/dataprovide_jalc

(6) アクセス数アラートサービス機能がある場合、登録希望者から設定希望があれば、設定する。

(7) 必要に応じ、作業内容を台帳(Excel 等)に記録する。例えば、以下のような内容を記載しておく、統計数算出等に役立つと思われる。

- 依頼日、登録完了日
- 依頼者氏名、依頼者の身分、所属(登録依頼時点)
- 登録依頼理由・動機等(分かれば)
- 研究データのタイトル
- 研究データのサイズ
- データのフォーマットタイプ
- 公開/非公開
- データ格納場所(リポジトリ内/外部)
- メタデータ URL(handle・DOI 等)

コラム「研究データ」登録の際に登録必須となる項目

公的資金によって生み出された研究データについては、「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※(必須項目)を含む項目が必要。(内閣府「研究 DX」参照)詳しくは「III 研究データのリポジトリ登録」を参照。

メタデータ共通項目をどのように JPCOAR スキーマで登録するかについては、pp.15-16 の図表を参考にすること。

※参考資料

1. [メタデータ共通項目](#)
2. [メタデータ説明書 1.0 版](#)
3. [メタデータ流通ガイドライン 研究データ編](#)
研究データの「メタデータ共通項目」と JPCOAR スキーマとのマッピングについて記載。
4. [メタデータ共通項目の出力 xml 例見本](#)
リンク先の 14_common_metadata_elements_cao.xml を参照。

※ 1、2 は内閣府のウェブサイト、3 は NDL のウェブサイト。(共通項目については内閣府、各項目のマッピングについては JPCOAR にお問合せ下さい)

特に研究データの場合、公募型の研究資金を受けて作成したにも関わらず、研究資金に関するメタデータ(助成情報、研究課題情報)が抜けていることが多いので積極的に確認したい。

助成情報、研究課題情報の特定・登録に於いては、当該研究者・研究室のみならず、研究推進室や外部資金部門との協力が推奨される。図書館にとっては正しい情報が入手でき、リポジトリで管理・登録されたデータは e-Rad や配分機関に実績報告をする際にも役立つ。

図:メタデータ共通項目を JPCOAR スキーマで入力した例

※JAIRO Cloud の場合: アイテムタイプは「デフォルトアイテムタイプ(フル)」を使用

作成者	情報
寄与者	
寄与者タイプ	HostingInstitution
寄与者識別子Scheme	ROR
寄与者識別子URI	https://ror.org/04ksd4g47
寄与者識別子	https://ror.org/04ksd4g47
姓名	国立情報学研究所
言語	ja
名前タイプ	Organizational
寄与者	
寄与者タイプ	DataManager
姓名	情報管理室
言語	ja
名前タイプ	Organizational
寄与者	
寄与者タイプ	ContactPerson
姓名	研究データテスト担当 contact@123456.jp
言語	ja
名前タイプ	Organizational
アクセス権	
アクセス権	open access
アクセス権URI	http://purl.org/coar/access_right/c_abf2
権利情報	
言語	en
権利情報Resource	https://creativecommons.org/licenses/by/4.0/deed.en
権利情報	Creative Commons Attribution 4.0 International
主題	
言語	ja
主題Scheme	e-Rad_field
主題	自然科学一般
内容記述	
内容記述タイプ	Abstract
内容記述	本アイテムは、研究データメタデータ共通項目登録のサンプルを作成したものです。
言語	ja
日付	
日付	2025-03-14
日付タイプ	Issued
資源タイプ	
資源タイプ識別子	http://purl.org/coar/resource_type/c_ddb1
資源タイプ	dataset
関連情報	
関連タイプ	isIdenticalTo
識別子タイプ	Local
言語	ja
関連名称	JPM1234567ABC-01
助成情報	
識別子タイプ	e-Rad_funder
助成機関識別子タイプURI	https://www.e-rad.go.jp/datasets/files/haibunkikan.csv
助成機関識別子	1234
助成機関名	〇〇機構
言語	ja
プログラム情報識別子タイプ	JGN_fundingStream
プログラム情報識別子タイプURI	MQQQ
言語	ja
プログラム情報	〇〇〇〇研究開発制度
研究課題番号	JPM1234567ABC
研究課題番号タイプ	JGN
研究課題名	××を用いた〇〇による〇〇の研究
言語	ja

表:メタデータ共通項目と対応する JPCOAR スキーマの一覧

番号	メタデータ共通項目	対応するJPCOARスキーマ
1	資金配分機関情報	助成機関識別子 (funderIdentifierTypeのみ) 助成機関名
2	体系的番号におけるプログラム情報コード	プログラム情報識別子
	プログラム名 (競争的研究費制度名称)	プログラム情報
3	体系的番号	研究課題番号
	プロジェクト名 (e-Rad課題名称等)	研究課題名
4	データNo.	関連情報
5	データの名称	タイトル
6	掲載日・掲載更新日 (メタデータ)	公開日 (日付_Issued / Updated)
7	データの説明	内容記述
8	データの分野 (e-Rad研究分野)	主題
9	データ種別	資源タイプ
10	概略データ量	サイズ (ファイル情報)
11	管理対象データの利活用・提供方針	権利情報
	アクセス権	アクセス権
	公開予定日	日付 (日付_Available)
12	リポジトリ情報	※対応無し (IRDBハーベスト時に補完)
	リポジトリURL・DOIリンク	識別子
13	データ作成者	作成者
	データ作成者のe-Rad研究者番号	作成者識別子
14	データ管理機関	寄与者姓名 (寄与者_HostingInstitution)
	データ管理機関コード	寄与者識別子 (寄与者_HostingInstitution)
	データ管理者	寄与者姓名 (寄与者_DataManager)
	データ管理者のe-Rad研究者番号	寄与者識別子 (寄与者_DataManager)
	データ管理者の連絡先	寄与者姓名 (寄与者_ContactPerson)
15	備考	※対応無し (非公開情報)

4. 登録完了後

4-1. 確認作業

以下の項目について確認を行う。

- ・ ランディングページについて
 - メタデータが正しく入力・表示されているか
 - (あれば)引用情報(JAIRO Cloud では Cite as の部分)が正しく表示されているか
- ・ 一次ファイルのダウンロード設定について
 - IP/PW 保持者のみダウンロード可となっているか
 - 公開予定日前にダウンロード不可となっているか
- ・ インデックス(カテゴリー)について
- ・ メタデータのハーベスト・連携について
 - IRDB に登録・修正内容が反映されているか
 - DOI や handle が有効になっているか

メタデータのハーベスト・連携については、反映・有効化されるまでに1週間程度時間がかかる場合もあるため、申請者へ連絡する際には、その旨を説明する。

必要に応じて以下のような各連携先へ接続し、検索してヒットするかどうか確認を行う。

連携先データベースの例: CiNii Research、NDL Search、Scopus、Web of Science、RePec、PubMed など

4-2. 依頼者への完了通知

- (1) 申請者へ登録完了の連絡をし、問題なくファイルがダウンロードできて開けるかどうか、確認してもらう。特にファイルのフォーマットなどの関係で、登録者側で確認できなかった点については、必ず確認してもらう。
- (2) 研究データの登録後に論文が公開される場合は、論文公開待ちリストを作成し、記録する。また、論文公開後に doi を連絡してもらうように依頼しておく。連絡がない場合には、リストを利用して、適宜督促や連絡を行う。科研費の実績報告書提出時、または年度末にまとめて行うなど、少なくとも年 1 回以上棚卸の機会を作っておくと良い。

コラム「データリポジトリにおける処理」

データに対する処理

ファイル形式とデータ変換

ファイル形式には、データ分析を行うことに向いているものと、長期保存に向いた形式がある。マイクロソフトの Word、Excel 形式、画像ファイルの jpeg 形式、動画ファイルの MPEG 形式、各種統計ソフトの独自のファイル形式は現状広く使用されているが、これらの仕様は公開されておらず、数 10 年後に使えなくなる可能性がある。この事態に備えて、仕様がオープンなファイル形式への変換を検討した方がよい。UK Data Service は[データの長期保存に適したファイル形式](#)をホームページ上で公開している。

また、データ変換時には文字コードにも注意する必要がある。一部のコードシステムは、すべての文字セットをカバーしていないので、UTF8 など汎用的な文字コードへの変換も検討した方がよい。

研究データを効率的に利活用できるようにするためには、ファイルの命名規則やフォルダ構成規則をリポジトリ内で決めておき、研究者から提供された研究データをその規則に従って変換することが望ましい。

ファイル名

ファイル名は、ファイルの内容、状態、バージョン等を含めて命名することで、ファイルの識別、分類、並び替えに役立てることができる。そのために、リポジトリ内の研究データについて命名規則を一貫させ、すべてのファイルに対して同一の命名規則を用いてファイルの命名を行っていることが望ましい。

ファイル命名規則を策定するにあたり、以下の点に気をつけるとよい。

- ・ 意味がありながら簡潔な名前をつけること
- ・ ファイル名を使用して、ファイルの種類进行分类すること
- ・ スペース、ドット、特殊文字（& または ? または !）の使用は避け、ファイル名の要素を区切るにはハイフン（-）またはアンダースコア（_）を使用すること。
- ・ 非常に長いファイル名は避けること
- ・ ファイル拡張子（例: .doc、.xls、.mov、.tif）は、ファイル名に含めないこと
- ・ ファイル名のバージョン番号を含めること

フォルダ構成

データファイルをどのように整理するかは研究者ごとに異なるが、リポジトリ内でフォルダ構造を規格化し、リポジトリでフォルダ構成を整えることは、ファイルやバージョンの検索と整理を容易にするために重要である。データフォルダには、データに関連するすべての資料を収納することが望ましい。

したがって、フォルダ構成として、研究データを収納したフォルダに並列して、データ収集方法に関する情報、データ収集用ツールに関する情報、収集したデータの処理に用いたツールに関する情報をそれぞれフォルダごとに収納することが望ましい。データを入手した利用者の利便性を考慮

し、フォルダの階層は 3 または 4 の深さとし、同階層に置くフォルダは 10 を超えないように整理すると望ましい。

データの再構築

変数名の再定義

データファイル内の変数名は、解析ソフトウェア操作の呼び出しコードとして利用されるので、以下のルールにしたがって命名されていることが望ましい

- ・ 変数は短く命名すること。標準では 8 文字を越えないようにすること
- ・ 解析ソフトウェアによっては国特有の文字は利用できないので、できるだけ避けること
- ・ 変数名 1 文字名に #、&、\$、@ などの特殊文字を利用することは避けること
- ・ 変数名に空白(スペース)を含まないこと、_(アンダーバー)等にする
- ・ 意味のある変数名にすること、アンケート調査データの場合、Q1a、Q1b、Q2 などと命名すると良い

図表の作成

研究データは、数値データやテキストデータの形式で記録されることが多いが、そのままでは直感的な理解が難しく、データの持つ価値が十分に活かされない可能性がある。そこで、グラフ、ヒートマップ、ネットワーク図、地図、統計表 などの可視化手法を活用することで、データのパターンや傾向を明確にし、再利用性を向上することが望ましい。

以下では、日本の研究機関の具体例について紹介する。

(1) 「ナショナルバイオリソースプロジェクト(NBRP)」

ナショナルバイオリソースプロジェクト(NBRP)は、文部科学省が支援する生物学研究データの収集・管理・提供を目的としたプロジェクトである。例えば、モデル生物の遺伝子情報や系統情報を 系統樹の図 や 遺伝子発現マップ として整理し、データベース上で公開している。これにより、国内外の研究者が視覚的にデータを確認し、比較検討できるようになっている。

(2) 「日本気象庁の気象データ可視化」

気象庁は、日本の気象データを 地図ベースの可視化 により公開している。例えば、温暖化傾向や降水量の変化をヒートマップや折れ線グラフで示し、一般市民から研究者まで幅広い層にとってデータが活用しやすいようになっている。これにより、環境研究や防災計画の立案に貢献している。

匿名処理

研究データの匿名処理とは、収集したデータから個人を特定できる情報を削除または変換し、データ主体のプライバシーを保護するプロセスを指します。日本における匿名処理は、「個人情報の保護に関する法律(個人情報保護法)」に基づいて行われる必要があります。特に、研究機関が扱うデータが個人情報に該当する場合、その適切な匿名化は法的義務として求められる。2022 年に改正された個人情報保護法において、「匿名加工情報」と「仮名加工情報」の 2 種類の匿名処理が定義されている。

匿名加工情報(法第 43 条)

個人情報加工し、特定の個人を識別できず、元の個人情報に戻せないようにしたデータ。研究データとして利用する場合、識別可能な情報(氏名、住所、電話番号など)の削除、識別符号の変更、結合可能な情報の分離などの手法が求められる。匿名加工情報を作成した事業者(または研究機関)は、加工方法や安全管理措置を公表する義務がある。

仮名加工情報(法第 41 条)

研究機関内でのデータ解析のために、氏名や連絡先などの識別情報を削除または置換することで、個人を直接識別できないようにしたデータ。ただし、一定の手続きを経れば元の情報に戻すことが可能であるため、外部提供には注意が必要である。研究目的での内部利用に適しており、個人情報保護の負担を軽減するための仕組みとして活用されている。

研究データの匿名化に際して、以下の方法が用いられている。

- ・ 直接識別情報の削除
氏名、住所、電話番号、メールアドレスなど
- ・ 間接識別情報の一般化
生年月日 → 年代、詳細な住所 → 市区町村レベル
- ・ データの集約・カテゴリ化
- ・ 年収 → 所得階層、学歴 → 高卒/大卒/大学院
- ・ 乱数やハッシュ化による符号化
ID や個人番号をランダムなコードに置き換える
- ・ データの部分的マスキング
例: 田中→T さん、03-xxxx-xxxx など

匿名化の程度は、研究の目的やデータの利用範囲に応じて慎重に決定する必要がある。また、研究データの匿名化は、個人情報保護法を遵守するだけでなく、研究倫理の観点からも適切に実施する必要がある。日本学術会議や各大学の倫理委員会では、研究データの適正な管理についてガイドラインを設けており、匿名処理の方法やデータの保存・利用について事前審査を受けることが推奨されている。また、研究対象者(被験者)に対し、匿名化の方法やデータの利用目的を事前に説明し、同意を得るプロセスも重要である。匿名化が不完全な場合や、再識別のリスクがある場合には、データの取り扱い方針を明示し、適切な管理体制を整えなければならない。

IV. 研究データの内容・メタデータ更新

研究データの登録完了後、

- ・ 別のファイルを追加する・別のファイルに更新する
 - 例)別のバージョンのデータが生成されたときに、別のファイルに差し替える
- ・ メタデータを修正・追記する
 - 例)根拠とした論文が公開されたときにメタデータを追記する

ということが考えられる。

特に、登録した研究データを根拠とした論文が公開された場合は、論文の DOI も関連リンクとしてメタデータに追記できるとよい。

1. 更新準備

(1) 登録データ・メタデータの再受領

ファイルを更新・追加したい場合はそのファイルを、メタデータを修正したい場合は修正後のメタデータを研究者から送付してもらう。

※ファイルを更新・追加する場合は、2. 更新作業で日付とバージョン情報を入力するため、ご連絡いただく必要がある。変更が軽微の場合や登録して間もない場合は更新しないこともある。申請者と要調整。

(2) 内容の確認

内容に不備がないか確認する。

※2.提出物の確認作業も参照。

2. 更新作業

ファイルの更新・またはメタデータの修正作業を行う。

ここでは、特にファイルの更新とメタデータの追記について説明する。

ファイルの更新

ファイルの更新時には、ファイルのバージョンと日付も変更する。

JPCOAR スキーマにおける、

- ・ ファイル情報
 - 日付

- バージョン情報

にあたるメタデータを修正する。

メタデータの追記

研究データの根拠となる論文の DOI を追加する場合は、関連情報に入力する。

JPCOAR スキーマにおける、関連情報の isSupplementTo 属性にあたるメタデータ項目に追記を行う。

3. 更新完了後

申請者に連絡する。

V. 研究データの削除・非公開

この章では、研究データを削除する、または非公開にするケースについて説明する。

1. 研究データの非公開

1-1. 概要

研究データの非公開は、将来的に再度公開する可能性があるデータに対して、一時的に公開停止の措置を取る行為である。一度公開した研究データを非公開にするケースとしては、公開制限があるファイルを誤って公開してしまった場合、登録者から希望があった場合、研究データの公開によって第三者の権利侵害が生じている場合などが考えられる。

1-2. 対応方法

A. 公開制限があるファイルを誤って公開してしまった場合

気が付いた時点で直ちに公開停止の措置を取り、影響範囲を調査する。平行して、適切な公開条件への修正を行う。

B. 登録者から希望があった場合

機関の取り下げポリシーを定めておき、登録者の意向と理由をもとに判断する。その際、期限付き非公開なのか削除なのかを明確にしておく。

C. 第三者の権利侵害が生じている恐れがある場合

登録者に連絡を取り、権利侵害への対応方法を協議する。緊急を要する場合は、機関リポジトリ管理者の権限で一時的に非公開措置を取る場合もある。

なお、研究データを非公開にした場合は、ランディングページに将来的な公開日時または研究データへのアクセス方法を記入しておく必要がある。具体的な記載例については、以下を参照のこと。

2. 研究データの削除

2-1. 概要

研究データの削除は、何らかの事情によって機関リポジトリ上からデータを削除する行為である。

一度公開した研究データを削除するケースとしては、

- ・ そもそも公開できない研究データであった場合

- ・ 登録者から希望があった場合
- ・ 重複登録があった場合
- ・ 所属機関のポリシーにより削除の決定がなされた場合

などが考えられる。

機関リポジトリ上で研究データを公開した場合、データ本体が不特定多数に提供されるだけでなく、IRDB などを通じてメタデータが広くウェブに流通することになる。また、削除した研究データがある論文の根拠データであった場合、当該論文に記載された内容との不整合が生じる可能性がある。公開の影響力を完全になくすことは非常に困難となるため、削除の判断は非公開よりも慎重に行うべきである。

2-2. 対応方法

A. そもそも公開できない研究データであった場合

契約やポリシーによって公開できず、また機関のストレージに保管することもできなかった場合は、気が付いた時点で直ちに公開停止の措置を取り、影響範囲を調査する。平行して、公開してしまった原因、理由を調査し、所属機関による合意のもとデータを削除する。

B. 登録者から希望があった場合

機関のポリシーを定めておき、登録者の意向と理由をもとに判断する。その際、期限付き非公開なのか削除なのかを明確にしておく。

C. 重複登録があった場合

データの流通状況やダウンロード件数など影響範囲を調査し、より影響が小さいほうを優先的に削除する。削除が難しい場合は、お互いのランディングページに「同一のデータである」旨を記述し、”isIdenticalTo”など関連情報を付与する。

D. その他、所属機関のポリシーにより削除の決定がなされた場合

削除に至った経緯などを整理し、機関内での合意プロセスに沿って処理する。

研究データに DOI を付与していた場合は、DOI の取り下げにあたり所定の手順を踏む必要がある。詳細は以下のガイドラインを参照されたい。

- ・ [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：JPCOAR スキーマ ver2.0.x 編](#)
- ・ [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：JPCOAR スキーマ ver1.0.x 編](#)
- ・ [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](#)

研究データを削除した場合であっても、機関リポジトリ上に墓標となるページを提供し、メタデー

タの公開を維持する必要がある¹。墓標ページには、以下のような内容を含めることが望ましい。

- ・ 削除日(例:2025 年 3 月 1 日)
- ・ 削除理由(例:データ提供者の要請による削除 (または プライバシー保護の観点から非公開化、倫理的配慮によりアクセス制限 など))
- ・ 関連情報(例:削除されたデータについては、新バージョン(DOI: 10.1234/...)が公開されています。)
- ・ 問い合わせ先(例:本データに関するお問い合わせは、データ提供者 (email@example.com)までお願いします。)

¹ <https://coar-repositories.org/coar-community-framework-for-good-practices-in-repositories/>

VI. 登録後の対応

1. 統計情報の取得

- ・ ダウンロード数、閲覧数、引用数などのデータを定期的に取得し、関係者へ共有する。共有先のフォーマットによって必要な情報が異なってくるため注意。
- ・ アクセス傾向の分析を行い、リポジトリ運営の改善に活かす。

登録パターン 1 の場合

資源タイプで絞り込むことで対象の統計情報を取得することができる。

登録パターン 2 の場合

ファイルタイプ別に抽出することで対象の統計情報を取得することができる。

2. 問い合わせ内容の記録管理

- ・ 研究者や利用者からの問い合わせを記録・管理し、FAQ を整備する。
- ・ よくある質問については、ガイドラインやリポジトリのウェブサイト上で公開すると良い。