

研究データを公開する

本コンテンツは「AI等の活用を推進する研究データエコシステム構築事業」における大阪大学の成果である研究データ管理教材「オープンサイエンス時代における研究データマネジメントの基礎について学ぶ」を基に、大阪大学・九州大学・オープンアクセスリポジトリ推進協会が連携して改修したものです。

研究データ公開の意義と動向

◆ 研究データ公開の意義

研究成果の透明性や公正性の確保

研究資金助成機関からの要求への対応

研究成果への直接的な引用や再利用によるインパクト

新たな研究コミュニティの創出や産学連携への発展

派生的な成果や新たな解釈への発展によるイノベーションの創出

◆ 研究データ公開の動向

● G7サミット(2023.5)における共同声明

G7広島首脳コミュニケ[2023/5/20 広島]

- G7は、FAIR原則に沿って、科学的知識並びに**研究データ及び学術出版物を含む公的資金による研究成果の公平な普及による、オープン・サイエンスを推進**する。

● 国内の施策

統合イノベーション戦略2023[2023/6/9 閣議決定]

- G7広島サミット・G7仙台科学技術大臣会合を踏まえ、我が国の競争的研究費制度における**2025年度新規公募分からの学術論文等の即時オープンアクセス**の実現に向けた国の方針を策定

公開前の確認事項

- ◆ 関係者からの許諾
許諾の文書化、ポリシー等の法的担保の有無
- ◆ 匿名化处理(個人情報への対応)
Cf., 総務省(2023, 6, 16)「匿名データの作成・提供に関するガイドライン」
https://www.soumu.go.jp/main_content/000806227.pdf
- ◆ データの粒度
再利用・引用に適した単位
- ◆ フォーマット
適切なファイルフォーマット
- ◆ 管理責任者(問合せ先)の決定
- ◆ エンバーゴ期間
「論文出版まで非公開」などの制約がないか
- ◆ データ管理計画(DMP)



Findable: 発見可能性

Accessible: アクセス可能性

Interoperable: 相互運用可能性

Reusable: 再利用可能性

- **FORCE11(The Future of Research)
Communications and e-Scholarship: IT活用による学術コミュニケーション促進に向けた研究者、図書館員、出版者、助成団体等のコミュニティ**
<https://www.force11.org/group/fairgroup/fairprinciples>
- **NBDC研究チーム(訳) FAIR原則(「THE FAIR DATA PRINCIPLES」和訳)(2019)**
<https://doi.org/10.18908/a.2019112601>

研究データの公開

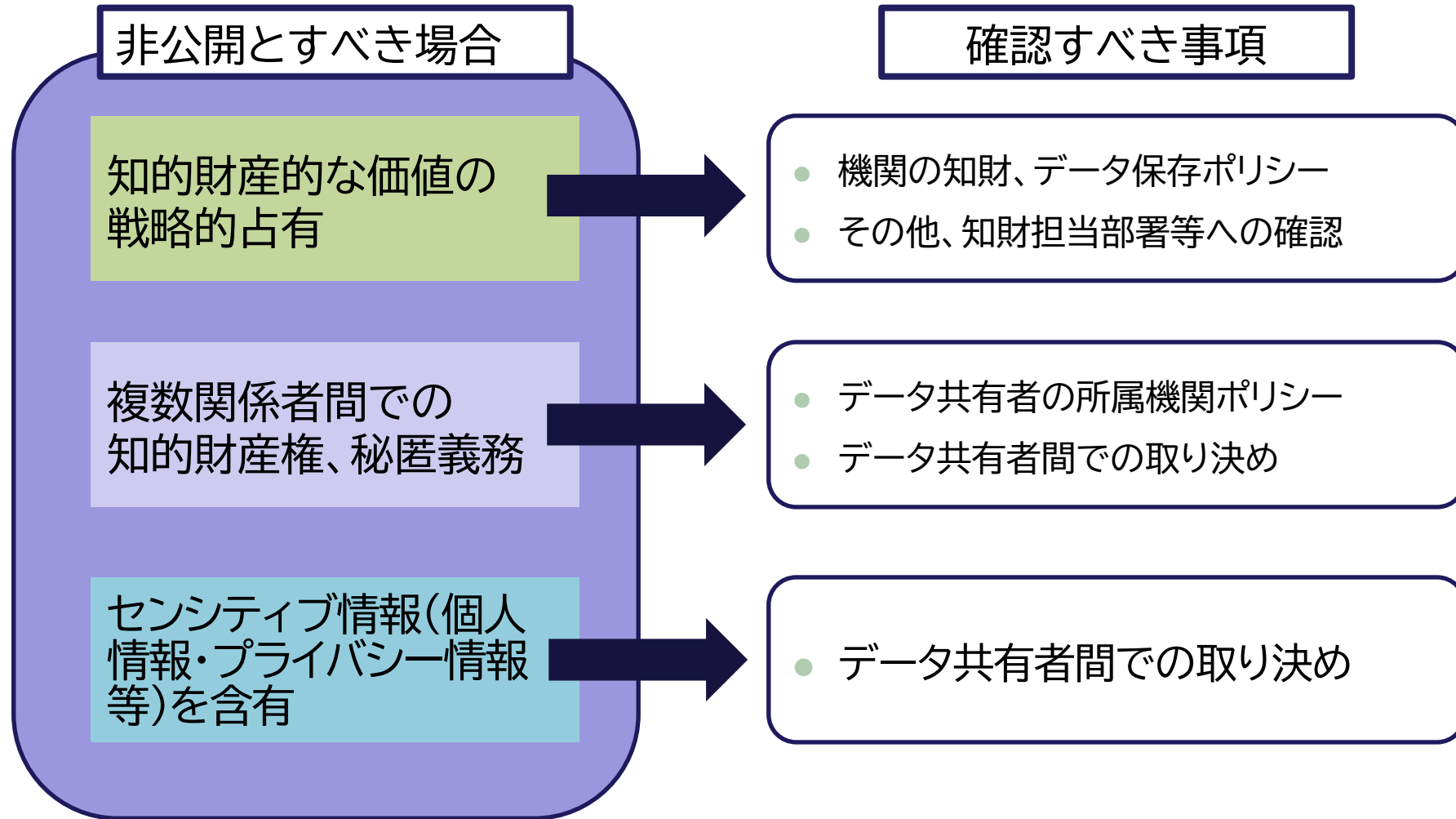
5.3 データセットの整備

データセットの整備

公開したデータを誰でも使えるようにするには、以下のような作業も必要です。

- 質の維持(クオリティコントロール)
 - データが整理されている、一貫性がある、エラーがないことが重要
- データに関する文書(3.3で説明)一式の作成と公開
 - データの概要を示した文書の作成
 - Readmeファイル
 - ユーザーガイド
 - 必要に応じて、他の人がデータを見たときに、データの作成経緯や、変数の説明などデータの構成等が分かる文書の作成
 - データに何らかのコードを付与している場合のマニュアル(codebook)
 - データを処理したプログラム
 - 質問紙調査

研究データの公開



研究データの公開

5.1 データ公開の意義と公開・非公開の区分

公開と非公開の区分

一般的に、研究データの公開及び非公開には以下の区分があり、取扱いには考慮が必要です。

1. **非公開(クローズ)**:個人での保管
2. **制限共有(セミクローズ)**:研究グループ等の構成員限りでの共有
3. **制限公開(セミオープン)**:限定された者への公開
4. **一般公開(オープン)**:制限のない公開

https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/036/attach/1378756.htm
第8期学術情報委員会(第8回) 配付資料
資料3 学術情報のオープン化に係る研究データの公開等について(案) より

単にオープン化を進めるのではなく、所属機関、研究者自身、研究コミュニティなどでオープン・アンド・クローズ戦略を十分に検討、適切な方法を採用する必要があります。

研究データ公開に関わるポリシー

◆ 機関レベルのデータ共有・公開の方針例

● 国立環境研究所

「国立研究開発法人国立環境研究所データの公開に関する基本方針(データポリシー)」(2017年4月1日)

https://www.nies.go.jp/kihon/kitei/kt_datapolicy.pdf

● 日本医療研究開発機構

「ゲノム医療実現のためのデータシェアリングポリシー」(2016年4月)

<https://www.amed.go.jp/content/000023353.pdf>

https://humandbs.biosciencedbc.jp/files/DAC/4th_meeting/2_ref_AMED_DSP.pdf

共同研究契約や個別の契約により公開制限が定められている場合は、研究パートナー、委託元などとの契約を遵守する必要がある

◆ 出版社のデータポリシー例

● データの共有ポリシーを定めている例

- Research Data Guidelines (Elsevier)
<https://www.elsevier.com/researcher/author/tools-and-resources/research-data/data-guidelines>

- Research Data Policy (Springer Nature)
<https://www.springernature.com/gp/authors/research-data-policy>

データポリシーが類型化されており、雑誌によりどの方針を採用しているかは異なる。

● データのアクセス先の明記を定めている例

- Data Availability Statements (Springer Nature)
<https://www.springernature.com/gp/authors/research-data-policy/data-availability-statements/>

データポリシーのType3, 4で、基礎データの利用可能性と利用方法について論文中に明記する「データ利用可能性ステートメント」を義務化

機関リポジトリとは

◆ 機関リポジトリとは

- 大学などの学術機関によって設置される、所属構成員による研究成果等を収集・保存・公開する電子アーカイブ
- 各機関の研究成果の可視性向上、社会に対しての説明責任を果たす役割

◆ 日本の機関リポジトリ

- 千葉大学(2005年)を皮切りに、現在は906機関*1 が運用(2023年3月時点)
- NIIがクラウドサービス(JAIRO Cloud)を提供
- 紀要論文や博士論文など学内刊行物がコンテンツの多くを占めるが、学術雑誌論文の公開に注力する機関もある。
- 研究データの公開先のひとつとしても期待されている*2

• 学術機関リポジトリ構築連携支援事業(2023, 4, 6)「機関リポジトリ公開数とコンテンツ数の推移」<https://www.nii.ac.jp/irp/archive/statistic/>
• 内閣府(2021, 6, 18)「統合イノベーション戦略2021」<https://www8.cao.go.jp/cstp/tougosenryaku/2021.html>

データリポジトリとは

一般に研究データやそれに付随するプログラム等を収集・保管の対象とするリポジトリ

◆ データリポジトリの種類

分野別リポジトリ

- re3data
分野を絞り検索可能
<https://www.re3data.org/>

機関リポジトリ

- 京都大学学術情報リポジトリ
<https://repository.kulib.kyoto-u.ac.jp/dspace/>
- 学術機関リポジトリデータベース
<https://irdb.nii.ac.jp/>

汎用リポジトリ

- Zenodo
<https://zenodo.org/>
- Figshare
<https://figshare.com/>
- Elsevier Mendeley Data
<https://data.mendeley.com/>

◆ 公開場所の選定の際に考慮すべき情報

出版者・助成機関の
ポリシー

機関の研究データ保
存ガイドライン

研究コミュニティ
へのインパクト

第三者による認証
(re3data で絞込検索可)

◆ 出版社の推奨するリポジトリ

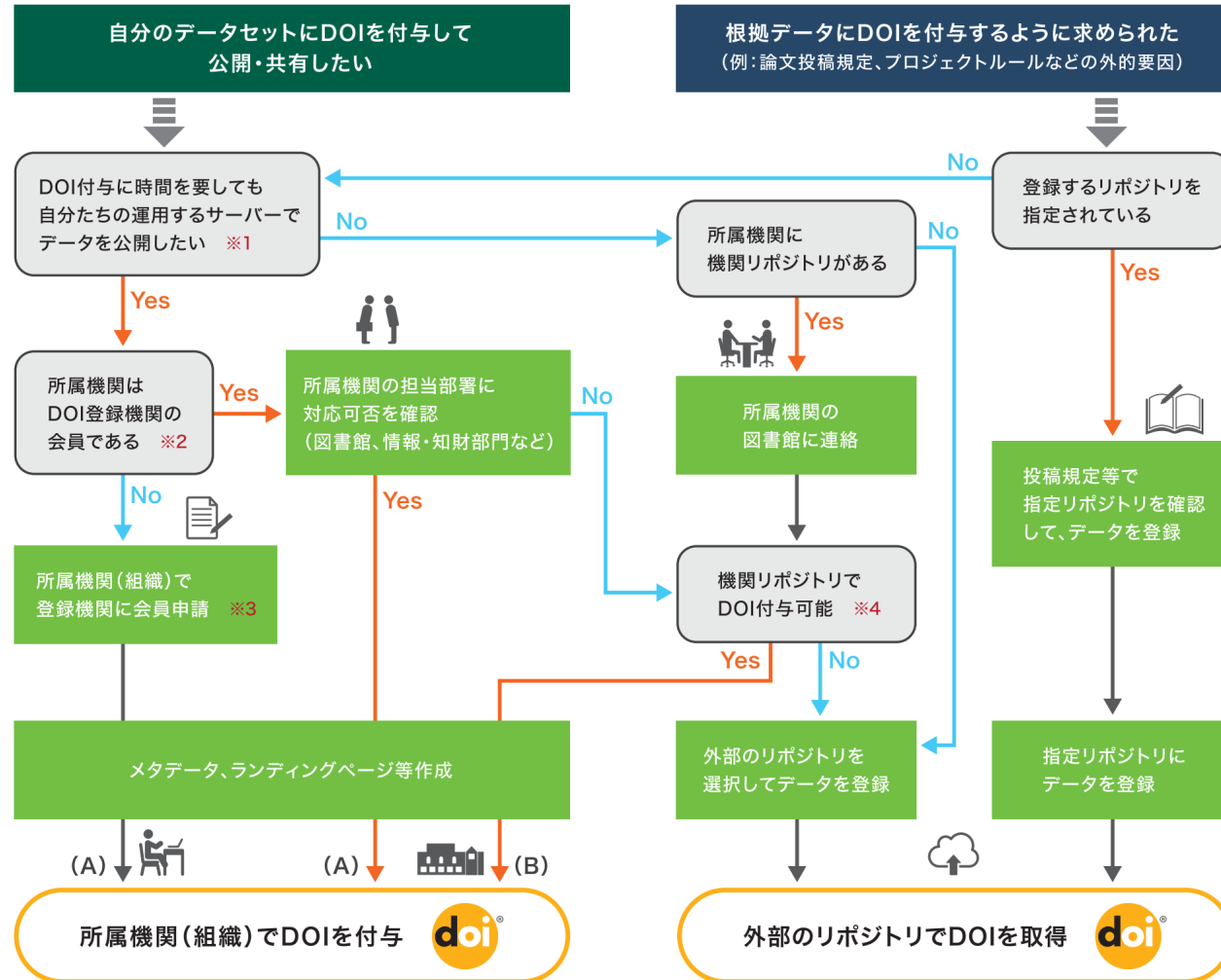
• Elsevier: Mendeley Data
<https://data.mendeley.com/>

• Springer Nature推奨リポジトリ一覧
<https://www.springernature.com/gp/authors/research-data-policy/recommended-repositories>

• PLOS ONE推奨リポジトリ一覧
<https://journals.plos.org/plosone/s/recommended-repositories>

研究データにDOIを付与する

研究データにDOIを付与するには？



メタデータとは

5.4 メタデータ

メタデータの定義と意義

- **メタデータ**とは
 - データに関するデータ
- 特徴:
 - データの詳細情報や文脈を記述する情報
 - データに関する文書は、人の利用を想定
メタデータは、特に機械処理を想定
- 意義:
 - 研究データの検索に必要な要素を提供
 - データを引用するための情報を提供
 - 将来のデータの再利用に備え、非公開の状態でも適切なメタデータを付与し、検索可能性や追跡可能性を担保しておくことが重要

メタデータとは

5.4 メタデータ

メタデータのタイプ

- 記述的メタデータ

- タイトル、著者、抄録、キーワードのようなユーザーが検索・閲覧を通してオンラインソースを発見するためのメタデータ
(利用者が**検索して、データにアクセスする**際に重要になる情報)

- 管理的メタデータ

- 保存や権利管理情報、フォーマットに関する技術的なメタデータ
(利用者が**実際にデータを利用する**際に重要になる情報)

- 構造的メタデータ

- データの構造やタイプ、データ間の関係を記述するためのメタデータ
(利用者が**データについて詳しく理解する**際に重要になる情報)

メタデータとは

5.4 メタデータ

さまざまなメタデータの標準

- 一般的なもの
 - Dublin Core (ダブリン・コア、略称: DC)
 - DataCite
- 分野特有のもの
 - Data Documentation Initiative(DDI, 社会科学)
 - DCCのwebサイトで検索可能

メタデータには、様々なタイプ・標準が存在します。研究データを公開し、検索されやすくするためには、同じ研究分野でよく使われているメタデータなど研究データに最も適したメタデータの標準を選択することが重要です。

ライセンスニング

ライセンスニング(利用許諾) = コンテンツの利用許諾の条件(ライセンス)を明示すること







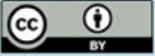
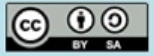
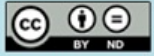


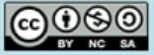
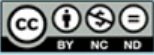
研究データの公開時に、メタデータにライセンスを記載しておくことで、一律な利用が可能になる

※ライセンスが付与されていなくても、その他の知的財産権で保護されている可能性がある点に留意

ライセンスリング

5.2 データの権利・ライセンス

ライセンスリング(2/2)

クリエイティブ・コモンズ・ライセンス	 表示 作品のクレジットを表示すること	 継承 元の作品と同じ組み合わせのCCライセンスで公開すること	 改変禁止 元の作品を改変しないこと
 表示 作品のクレジットを表示すること	 CC BY 表示	 CC BY-SA 表示-継承	 CC BY-ND 表示-改変禁止
 非営利 営利目的での利用をしないこと	 CC BY-NC 表示-非営利	 CC BY-NC-SA 表示-非営利-継承	 CC BY-NC-ND 表示-非営利-改変禁止

クリエイティブ・コモンズ・ライセンスを使用すると、表示・非営利・継承・改変禁止の条件を組み合わせ、利用範囲を表現できます。

データが著作物にあたらない場合は、「CC0」ライセンス(いかなる権利も保有しない)を付与します。

クリエイティブ・コモンズ作者のクレジットを表記することを求める CC BY をベースとして、非営利 (NC)、継承 (SA)、改変禁止 (ND) を組み合わせた 6 種類のライセンスが存在する。

長谷川世一「日本におけるクリエイティブ・コモンズ・ライセンスの実際」『情報管理』vol.58, no.5, p.344
<https://doi.org/10.1241/johokanri.58.343>

7

著者によるライセンス表示

Creative Commons License (クリエイティブコモンズ・ライセンス)

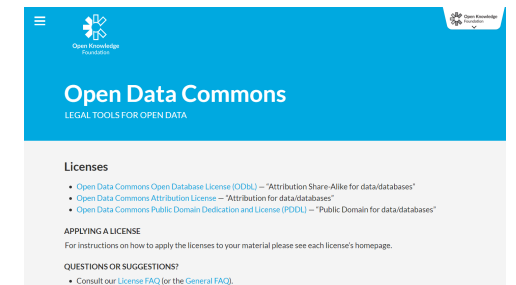
表示・非営利・改変禁止・継承の4条件を組み合わせて利用条件を表現するツール



<https://creativecommons.jp/licenses/>

Open Data Commons (オープンデータコモンズ)

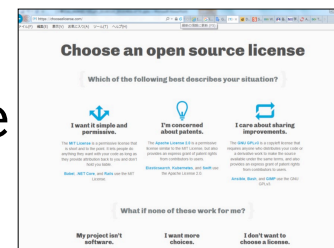
事実情報由来のデータ・データベースにも適用可能。



<https://opendatacommons.org/licenses/>

参考:ソフトウェアのライセンスツール
GitHub, Choose an open source license

<https://choosealicense.com/>



登録先のリポジトリによって、使えるライセンスは異なるため注意

管理者側からの意思表示

Public Domain Mark(パブリックドメインマーク)

- Creative Commonsによるツール
- 著作物を初めとする知的創作物について、知的財産権が発生していない状態を表す

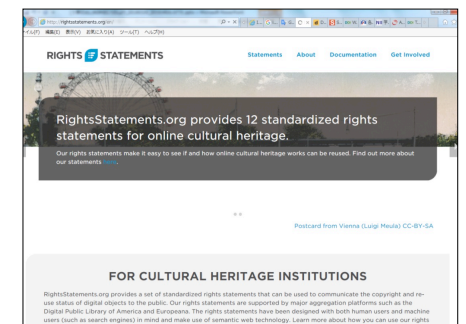


参考:Creative Commons, CC0 PDM comparison chart

https://wiki.creativecommons.org/wiki/CC0_PDM_comparison_chart

Rights Statements(ライツステートメント)

- 米国デジタル公共図書館(DPLA)とEuropeanaが2016年4月に創設
- オンライン文化資源の著作権のステータス交換が目的
- 著作権の状態を、「著作権が存在する場合」、「著作権が存在しない場合」、「著作権の存在が明確でない場合」に分類し、12種類のライセンスで表現



<http://rightsstatements.org/en/>

データ論文とは

◆ データ論文(データペーパー)とは

- 実験や観測で得られた数値などの「データ」につき、その取得方法や処理方法、所在等について記載した論文。通常の学術論文と異なり、新たな科学的知見が含まれているかどうかは問わない。

◆ データ論文を出版することによって

データの利用者拡大 検証可能性の向上 データ作成者、データ取得者の貢献度をより明確に表現

◆ データ論文を専門に受け付ける「データジャーナル」も存在

出版社のデータジャーナル	Scientific Data	https://www.nature.com/sdata/
	Data in Brief	https://www.sciencedirect.com/journal/data-in-brief
分野別のデータジャーナル	Geoscience Data Journal	http://rmets.onlinelibrary.wiley.com/hub/journal/10.1002/(ISSN)2049-6060/
	Research Data Journal for the Humanities and Social Sciences	https://brill.com/view/journals/rdj/rdj-overview.xml
日本のデータジャーナル	Polar Data Journal	https://pdr.repo.nii.ac.jp/

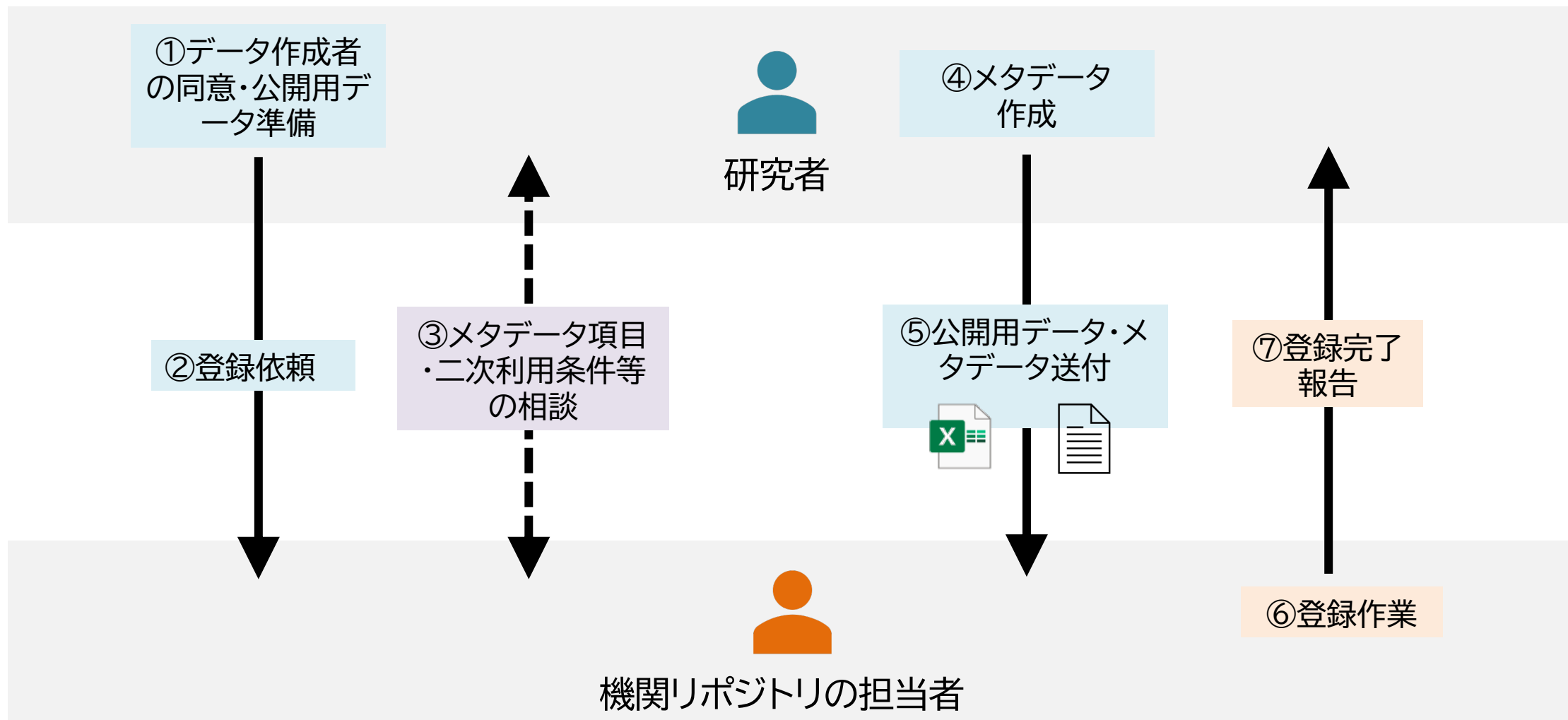
研究データ公開の流れ(機関リポジトリ)

◆機関リポジトリでの公開手順(一例)

- ① 研究データの作成に関わった者の同意を得る
- ② 公開する研究データファイルを準備する
- ③ 機関リポジトリの担当者にデータの公開希望を連絡する
 - 具体的な手順について担当者にご確認ください
 - データの種類、どのようなメタデータ情報を表示したいか、制限公開、二次利用ライセンスなどについて希望を伝えるとよいでしょう
- ④ 手順に従い、研究データ・メタデータを図書館に送付する

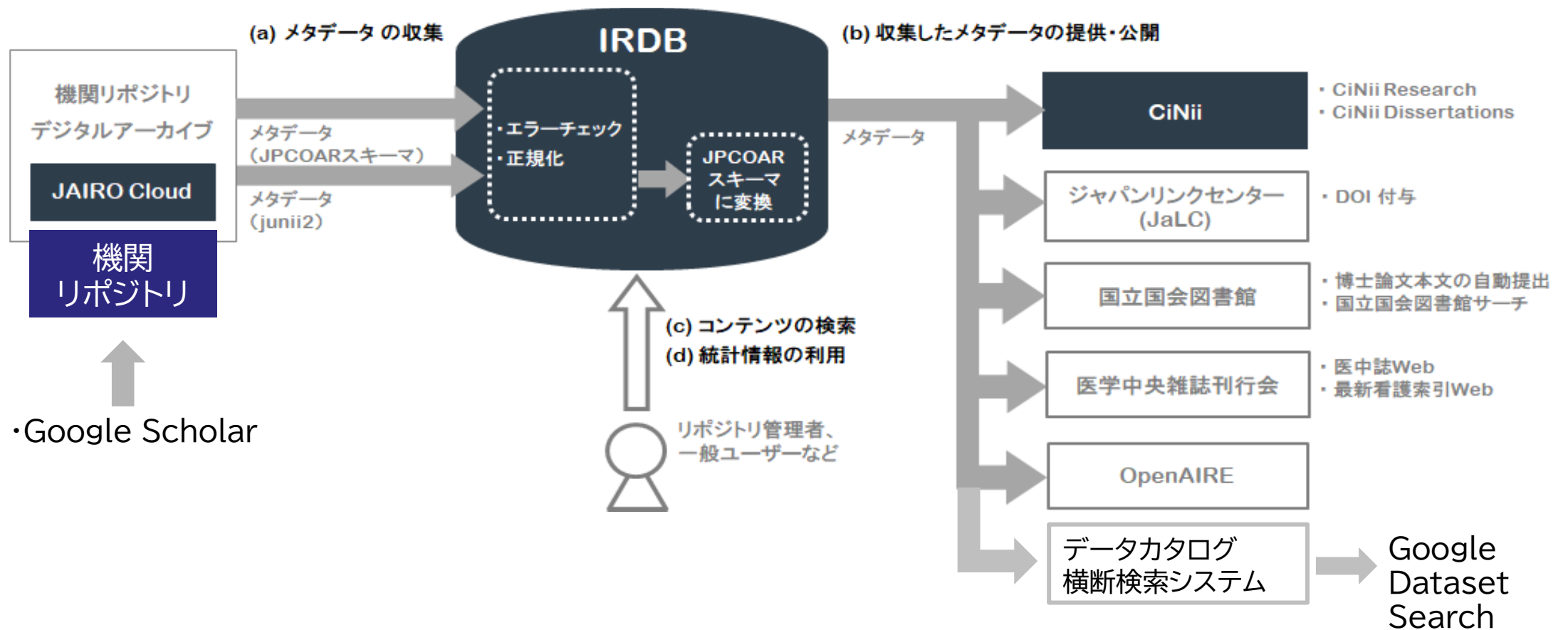
機関リポジトリで研究データを公開する

機関リポジトリでの公開手順図(一例)



メタデータの流通

メタデータはNII(国立情報学研究所)のIRDBというデータベースによって収集され各データベース等に流通する。



参照資料

- 学術機関リポジトリ構築連携支援事業(2023, 4, 6)「機関リポジトリ公開数とコンテンツ数の推移」
<https://www.nii.ac.jp/irp/archive/statistic/>
- 九州大学データ駆動イノベーション推進本部研究データ管理支援部門, 九州大学附属図書館図書館DX支援室(2023, 08)「はじめよう、研究データ管理」, 5_研究データの公開_202407改訂,
<https://doi.org/10.15017/6796147>
- 研究データ利活用協議会・リサーチデータサイテーション小委員会(2019, 12, 20)「研究データにDOIを付与するには？5分で分かる研究データDOI付与」. https://doi.org/10.11502/rduf_rdc_doileaflet
- 内閣府(2021, 6, 18)「統合イノベーション戦略2021」
<https://www8.cao.go.jp/cstp/tougosenryaku/2021.html>
- 西園由依 et al. (2020, 10, 29)「研究者のための研究データマネジメント_研究後_データ論文を通じたデータ公開」『教材「研究者のためのデータマネジメント」』. <https://jpcoar.repo.nii.ac.jp/records/294>
- 吉田幸苗 et al.(2021, 2, 10)「研究データ管理サービスの設計と実践_第5章_第2版」.
<https://jpcoar.repo.nii.ac.jp/records/607>