

RDMkit-jpを通じた 国内情報共有サービスの開発と展望

古川雅子1、南山泰之1、大波純一2、増井誠生1、長岡千香子1

(国立情報学研究所1、理化学研究所2)



概要

2021年の内閣府による「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」公開以降、国内で研究データをオープンにする取り組みが進んでいる。しかし、RDM(研究データ管理)の知識はまだ模索中である。一方、欧州ではFAIR原則に基づき、ELIXIRがRDMkitサイトを構築しており、これは国内でも有用と考えられる。そこで、我々は2023年に「RDMkit-jp」を公開した。本サイトではRDMに関する知識やツール情報を随時追加しており、GitHubを利用したフィードバックや、教材の簡易アクセスも開発中である。今後もコンテンツを充実させ、国内のRDM整備に資するサービスとしていく。



国内の研究データ管理(RDM)の課題

2022年以降、「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」のための「メタデータ共通項目」がガイドラインとして公開され、継続して更新されている。

内閣府 内閣府の政策 ~組織・制度 ~広報・報道 ~活動・白書等 ~情報提供 内閣府ホーム > 内閣府の政策 > 科学技術・イノベーション > 研究DX 研究DX(デジタル・トランスフォーメーション) 中略 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 → 本文(PDF形式:84KB) ** > 概要(PDF形式:935KB) ** >メタデータの共通項目 最新版(PDF形式:830KB) ™ ● 研究データ2022(PDF形式: 3289KB) *** 分割版1(PDF形式:1524KB) 1 2(PDF形式:1724KB) 1 3(PDF形式:1964KB 1 4 ● 研究データの管理・利活用に関する取組状況の評価体系への導入について(PDF形式:293KB) ● 研究データリポジトリ整備・運用ガイドライン(PDF形式:893KB) 動 ⑤国立研究開発法人におけるデータポリシー策定のためのガイドライン > 本文(PDF形式: 755KB) 動 > 解説(PDF形式:574KB) m

内閣府Webサイト

「研究DX - 科学技術・イノベーション」 https://www8.cao.go.jp/cstp/kenkyudx.html 大学や研究所などの公的機関は、内閣府のこれらの指針に対応し、研究データ管理体制を整備する必要がある。

 \downarrow

しかし国内の公的機関では、研究データ管理 (RDM; Research Data Management) の基礎知識やベストプラクティスなどの情報共有が十分ではなく、対応が困難。

基本的な考え方

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

> 令和3年4月27日 統合イノベーション戦略推進会議

1. 目的

デジタル・トランスフォーメーション (DX) の進展に伴い、大量のデータ、高度な計算資源、大規模なネットワークの活用など、研究開発活動の変容が進んでいる。また、研究成果の発表手段の多様化により、オープンサイエンス等の世界的な知の共有を目指した研究成果のオープン化が進みつつある。一方、世界的な出版柱による論文発表の寡占が進展するなど、研究成果や研究プロセス全体で得られたデータをビジネスの対象とする動きも見られる。研究データは、我が国のみならず世界にとって重要な知的資産であるといえる。このような状況を踏まえ、知の結合と発展を促し、優れた研究成果とイノベーションを創出していくためには、研究活動における自由と多様性を尊重しつつ、国際的な貢献と国益の双方を考慮に入れた、オープン・アンド・クローズ戦略に基づく研究データの管理・利活用を実行することが我が国として求められる。

第6期科学技術・イノベーション基本計画(令和3年3月26日閣議決定)においても、 第2章2(2)「新たな研究システムの構築(オープンサイエンスとデータ駆動型研究等の推 進)」において、公的資金による研究データの管理・利活用の推進について示されており、 このような我が国における基本的な考え方を具体化するため「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」(以下「本考え方」という。)を示す。

2. 研究データの定義等

2-1. 研究データ及び管理対象データの定義・範囲

推奨されるメタデータ共通項目

	「公的資金によ	る研究データ	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 令和6年7月30日 改正 関係所省中合せ
_	項目	必須/任意	備考
1	資金配分機関情報		公募型の研究資金を配分した資金配分機関 (府省含む) の名称 (英語略称可) 公募型の研究資金以外の場合は記入不要
	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 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		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
3	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	管理対象データを特定するための番号
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	メタデータを公開した日、更新した日
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)
9	データ種別	必須	「データセット」等を指定する。
LO	概略データ量	任意	管理対象データの概ねのデータ容量、またはシステムから出力される値
	管理対象データの利活用・提供方針	必須	ライセンス情報等の利用条件や制約条件を記載
	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
10	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
LZ	リポジトリURL・DOIリンク	任意	情報があれば記載
12	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
13	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
14	データ管理者	必須	データ管理機関において各管理対象データを管理する部署名または担当者の名前
	データ管理者のe-Rad研究者番号		研究者番号がない管理者や、管理者が組織の場合は不要。e-Rad研究者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



内閣府の指針に対応し、RDMを実施するための情報基盤の必要性

■国立国会図書館 「メタデータ流通ガイドライン:研究データ編」 https://iss.ndl.go.jp/information/guideline_researchdata/



■JPCOAR(オープンアクセスリポジトリ推進協会) セミナー資料

https://jpcoar.repo.nii.ac.jp/news/2022#news_20230309

2023年3月9日 第18回月刊JPCOAR「研究データのメタデータの共通項目への対応に向けて」の資料公開

3月14日(火)14:00~16:00開催の第18回学術コミュニケーションセミナー(月刊JPCOAR)「研究データのメタデータの共通項目への対応に向けて」の配布資料を以下のとおり事前公開しました。

大谷祐子(内閣府 科学技術・イノベーション推進事務局政策調査員(統合戦略担当))

講義:研究データのメタデータの共通項目の解説

鹿沼智美(国立情報学研究所学術基盤推進部学術コンテンツ課)

講義:メタデータ流通ガイドライン(研究データ編)の紹介

岩井雅史(信州大学附属図書館/研究データ作業部会・コンテンツ流通促進作業部会)

講義:メタデータの共通項目とJPCOARスキーマ

三上絢子(北海道大学附属図書館/研究データ作業部会・コンテンツ流通促進作業部会)

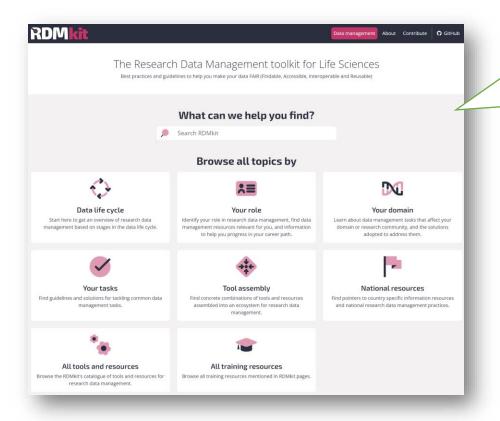
事例:学術機関での研究データ対応の事例紹介



RDMkit

■ RDMkit

https://rdmkit.elixir-europe.org/ 欧州ELIXIRが提供する生命科学研究データの RDMに関するトレーニングツール



立場や目的ごとの

- ・どのようなシナリオに従ってデータ管理するべきか
- ・最初に手をつけるべきデータ管理プロセスはどこか
- ・よくある課題
- ・関連ツール の情報が、**自然文でまとめられている**。

CC-BYライセンスで公開されているRDMkitの枠組みを利用し、国内のRDMを扱う立場や目的ごとに平易な言葉で情報共有できるための基盤を構築



トップページ (https://rdmkit.rcos.nii.ac.jp/)



2023年6月公開

以下16ページを作成し、研究データのライフサイクルや区分から辿れるように配置。

#	分類	ページ名
1	役割	研究者
2		データ管理者:研究
3		データ管理者:ポリシー
4		データ管理者:インフラ
5	タスク	識別子
6		既存データ
7		文書化とメタデータ
8		ライセンス
9		データ管理計画
10		データ整理
11		データ品質
12		データ分析
13		データ公開
14		データストレージ
15		センシティブデータ
16		コンプライアンスの監視および測定



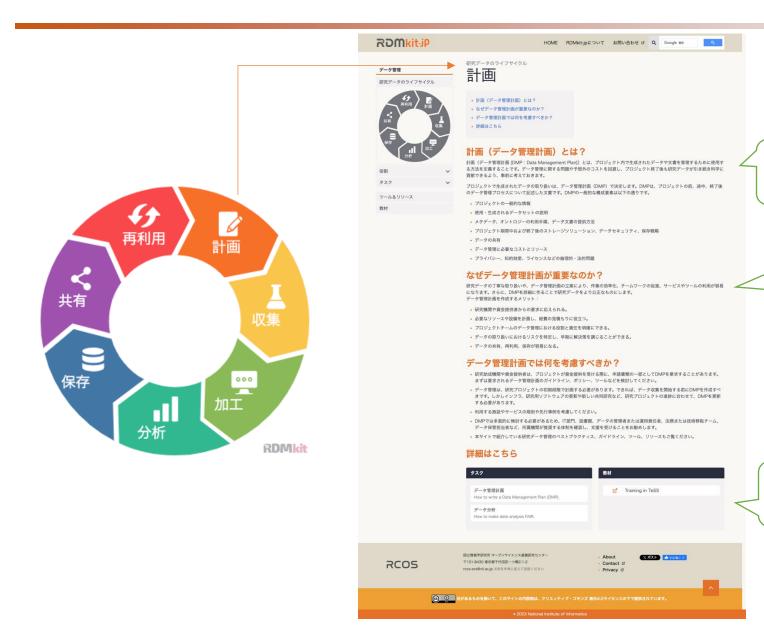
「研究データ管理」に関する 平易な解説

ライフサイクルの図形から 直接説明ページへ移動

ツール&リソース、教材については、 列挙してページに表示



体系化された内部のコンテンツ



RDM関連語句の平易な 日本語での説明ページ

理由付けや考え方を、利用者の立場に立って説明

言及されたリソースや 関連項目をリンク



体系化された内部のコンテンツ

12種のタスクに分けて、 ベストプラクティスとなる 例やノウハウを記述

- 1. 識別子
- 2. 既存データ
- 3. 文書化とメタデータ
- 4. ライセンス
- 5. データ管理計画
- 6. データ整理
- 7. データ品質
- 8. データ分析
- 9. データ公開
- 10.データストレージ
- 11.センシティブデータ
- 12.コンプライアンスの監視および測定



欧州版RDMkitの内容を 踏襲しつつ、日本国内のツー ルや考え方に置き換えて表現



体系化された内部のコンテンツ

4つに分類された役割から 対応シナリオやRDMに おける行動内容を説明

• 研究者

• データ管理者:研究

• データ管理者:ポリシー

データ管理者:インフラ



リソースリスト

NII RCOS

ROMkit-iP

HOME RDMkit-jp

言及されているツールやリソースのリスト。日本国内の情報も追記し、96件が登録されている。国内のバイオデータベースについては、Integbioデータベースカタログの説明文を参照し、作文作業にはChatGPTの提案を利用した。

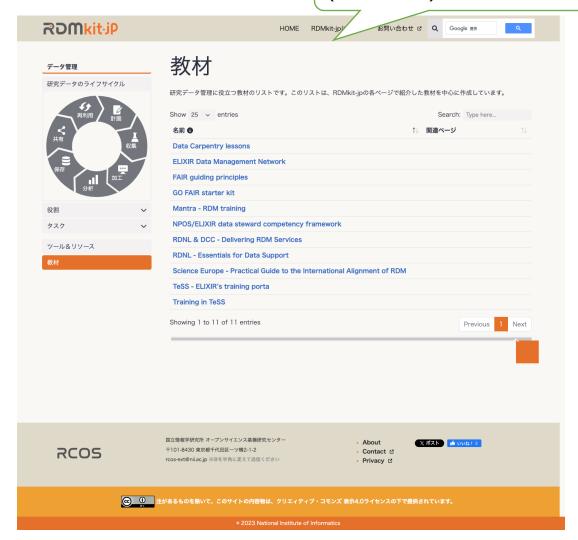


ツール&リソース

研究データ管理に役立つツールとリソースのリストです。ここに列挙したツールとリソースは網羅的なものではありません。本サイトの各ページでは、下記のテーブルをフィルターにかけたリストがページの最後に表示されます。ツールやリソースのリンクが切れている、または間違いがある場合は、「お問い合わせ」ページからお知らせください。

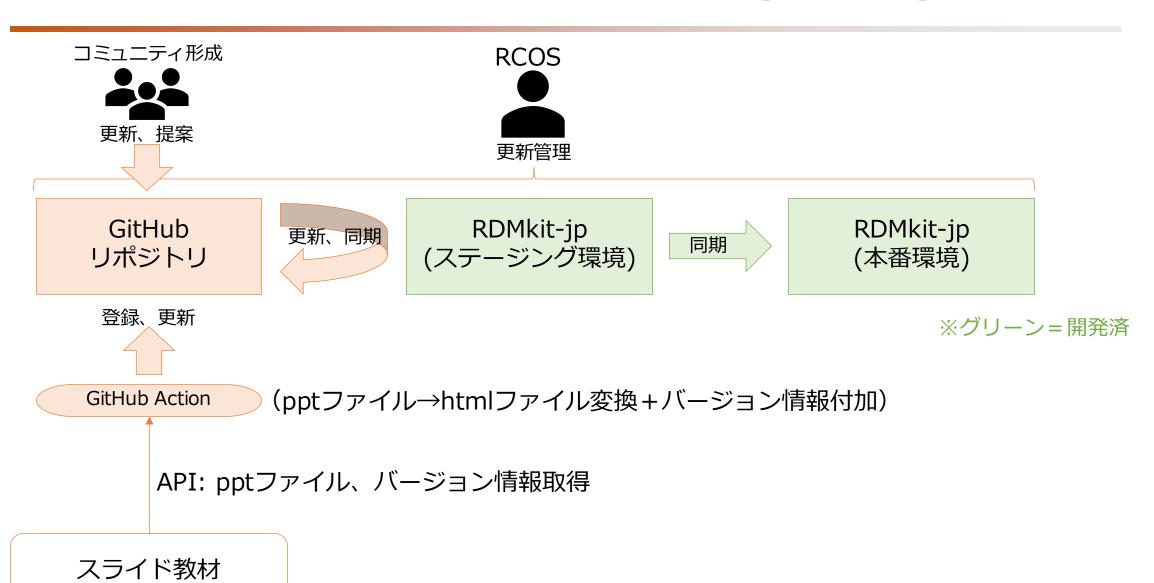


言及されている教材のリスト (更新準備中)





GitHubを利用した情報更新体制 (開発中)





まとめ・今後の展望

- 2021年に「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」が公開され、国内で研究データのオープン化が進展するものの、RDM(研究データ管理)の知識はまだ模索段階
- 欧州ではFAIR原則に基づき、ELIXIRがRDMkitサイトを構築
- 2023年に「RDMkit-jp」を公開し、RDMに関する知識やツール情報を随時追加
- GitHubを活用したフィードバック機能の導入、教材の簡易アクセス機能を開発中
- 今後、コンテンツの充実を図り、国内のRDM整備に貢献するサービスを目指す