

RDMkit-jp: 研究データ管理(RDM)の知識や情報を共有するサービス

○大波純一¹、増井誠生²、長岡千香子²、古川雅子²、南山泰之²

1.国立研究開発法人 理化学研究所 バイオリソース研究センター 遺伝子材料開発室、2.国立情報学研究所 オープンサイエンス基盤研究センター

要旨

2021年に内閣府から公開された「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」を皮切りに、国内の学術研究機関では、研究者が情報をオープンにするためのリポジトリやガイドラインの整備が進められている。これに必要なRDM (Research Data Management) の知識は研究コミュニティや大学図書館の現場で模索されている一方、欧州ではFAIR原則に基づいた基盤の整備が比較的早期に進められていた。例えば、生命科学の領域では政府間組織ELIXIRがツールや環境を整備しており、研究データ共有のための国際会議であるResearch Data Allianceとの連携を補助するRDA TIGERプロジェクトが設立されるなど、環境整備と実務の両面から成果を上げている。特にELIXIRが構築したRDMkit-jpは、RDMの知識を職種や研究データのライフサイクルごとに調べることができ、国内の学術機関としても有用だと考えられた。そのため、我々は2023年に「RDMkit-jp」を新しく構築し、国立情報学研究所より公開した。サイト内にはRDMI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機能も開発中である。今後も研究コミュニティとの連携や、フリーで利用できるコンテンツの充実をはかり、国内のRDM整備に資するサービスとしていく。

内閣府Webサイトから公開されている研究データ管理の指針

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。

RDMkit-jpの構成

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



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出されるように配置。

No.	名称	ページ
1	役割	研究者
2	データ管理	研究
3	データ管理	ポリシー
4	データ管理	インフラ
5	メタ	データ
6	保存	データ
7	次世代とデータ	データ
8	ライセンス	データ
9	データ管理	管理
10	データ	管理
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	データ	セキュリティ
16	コンテンツ	メタデータの整理および管理



「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

役割

タスク

ツール&リソース

教材

ツール&リソース、教材について、一覧表示

内閣府Webサイト「研究DX - 科学技術・イノベーション」
<https://www.8.cao.go.jp/cstp/kenkyudx.html>

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

1. 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

2. 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基づくメタデータの共通項目

基本的な考え方

1. 目的
デジタルトランスフォーメーション(以下「DX」)の進展に伴い、大量のデータが蓄積され、高度な分析・加工が求められる。研究機関は、研究活動の効率化を図るため、蓄積されたデータの活用を促進し、データの価値を最大化することを目指す。DXの推進には、データの管理・利活用に関する基本的な考え方を示すことが重要である。本指針は、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を示すことにより、研究機関のDX推進に資することを目的とする。

2. 原則
本指針は、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を示すことにより、研究機関のDX推進に資することを目的とする。本指針は、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を示すことにより、研究機関のDX推進に資することを目的とする。

体系化された内部のコンテンツ

プランニング

RDM関連語句の平易な日本語での説明ページ

データマネジメント・プランニングとは?

データマネジメント・プランニングとは、研究データのライフサイクル全体を通じて、データの品質、セキュリティ、アクセス性を確保するための計画を立てることを指す。これは、データの生成、保存、共有、廃棄の各段階において、適切な管理を行うことを意味する。

理由付けや考え方を、利用者の立場から説明

なぜデータマネジメント・プランニングが重要なのか?

データマネジメント・プランニングは、研究データのライフサイクル全体を通じて、データの品質、セキュリティ、アクセス性を確保するための計画を立てることを指す。これは、データの生成、保存、共有、廃棄の各段階において、適切な管理を行うことを意味する。

データマネジメント・プランニングでは何を考慮すべきか?

データマネジメント・プランニングでは、データの品質、セキュリティ、アクセス性を確保するための計画を立てることを指す。これは、データの生成、保存、共有、廃棄の各段階において、適切な管理を行うことを意味する。

詳細はこちら

言及されたリソースや関連項目をリンク

データマネジメントプラン

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

研究者

4つに分類された役割から対応シナリオやRDMIにおける行動内容を説明

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表現

推奨されるメタデータ共通項目

項目	必須	推奨
1. 識別	必須	必須
2. 説明	必須	必須
3. 作成日	必須	必須
4. 作成者	必須	必須
5. 更新日	必須	必須
6. 更新者	必須	必須
7. 権利	必須	必須
8. 利用	必須	必須
9. 保存	必須	必須
10. 共有	必須	必須
11. 廃棄	必須	必須
12. 評価	必須	必須
13. 監査	必須	必須
14. 変更履歴	必須	必須
15. 関係性	必須	必須
16. 参照	必須	必須
17. 注釈	必須	必須
18. 問い合わせ先	必須	必須
19. 関連情報	必須	必須
20. その他	必須	必須

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ります。

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。

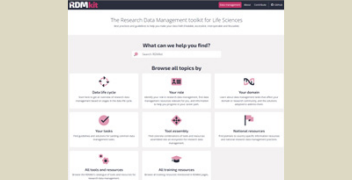
国内の研究データ管理(RDM)の課題

内閣府の指針に対応しRDMを実施するための「知識」を共有

■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/



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



■JPCOAR(オープンアクセシ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



立場や目的ごとの、どのようなシナリオによってデータ管理すべきか、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、平易な言葉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

リソースリスト

ツール&リソース

言及されているツールやリソースのリスト。日本国内の情報を追記し、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、Integrio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の提案を参考とした。

教材

他のページから言及されている教材のリスト

まとめ・今後の展望

- 公的資金による研究データの管理・利活用が必要とされる背景
- 2023年6月に、RDMの基礎情報や公開例を体系化したサイトをリリース
- 国内のRDMIに資する学術基盤へのドキュメントやサポート
- 本家RDMkitは欧州とバイオ分野に特化した内容だったが、日本語に書き換えローカライズ
- 内容については、今後フィードバックを受け品質向上を継続し、本家との連携も検討
- 国内のRDM情報を必要とする分野との協調体制のた、協力者を募集中
- GitHubリポジトリとシステム連携し、利用者Issueのフィードバックや版管理に利用予定

謝辞

本研究は文部科学省「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」の支援を受けたものです。