

RDM知識情報の共有サイト



大波純一¹、塩谷昌之²、南山泰之²、増井誠生³、長岡千香子³、古川雅子³
(理化学研究所¹、東京大学社会科学研究所²、国立情報学研究所³)

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。

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。



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。

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>



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) 無
- 概要(PDF形式:935KB) 無
- スターデータの共通項目 最新版(PDF形式:830KB) 無
- 研究データ2022(PDF形式:3289KB) 無
- 別冊版1(PDF形式:1524KB) 無 2(PDF形式:1724KB) 無 3(PDF形式:1964KB) 無 4(PDF形式:2218KB) 無
- 研究データの管理・利活用に関する取組状況の経緯体系への導入について(PDF形式:293KB) 無
- 研究データリポジトリ登録・運用ガイドライン(PDF形式:893KB) 無
- 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) 無
- 解説(PDF形式:574KB) 無

基本的な考え方

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、研究開発活動の姿容が進んでいる。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す研究開発のオープン化が進みつつある。一方、世界的な出版社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、知の結集と発展を促し、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、国際的な貢献と公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

関係府省庁承認

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
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
4 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番号 公募型の研究資金以外の場合は記入不要
5 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
6 データNo.	必須	管理対象データを特定するための番号
7 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
8 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
9 データの説明	必須	端的かつ中身の分かる内容を記載
10 データの分野	必須※	e-Radの研究分野(主分野)
11 データ種別	必須	「データセット」等を指定する。
12 総データ量	任意	管理対象データの総データ容量、またはシステムから出力される値
13 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
14 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
15 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
16 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
17 リポジトリURL・DOIリンク	任意	情報があれば記載
18 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
19 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
20 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
21 データ管理機関コード	任意	データ管理機関のコード
22 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名
23 データ管理者のe-Rad研究者番号	必須	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
24 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
25 備考	任意	

※公募型の研究資金による研究活動の場合

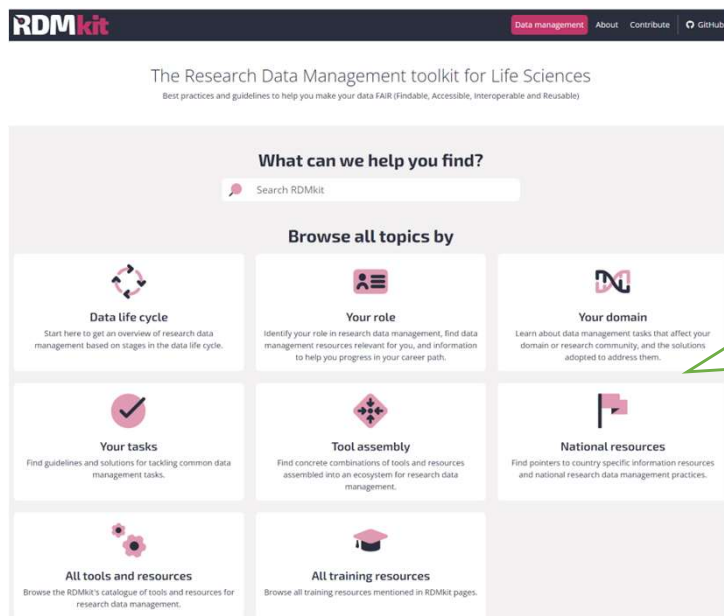
RDM情報公開サイトの状況

国内外の分野・組織ごとに多様なRDMの説明を独自に作成している状況

No	ウェブサイト名称	作成母体	国	分野	URL	備考
1	RDMkit	Elixir	EU	生命科学	https://rdmkit.elixir-europe.org/	
2	UK Data Service	UKRI	UK	社会科学	https://ukdataservice.ac.uk/	
3	ICPSR	University of Michigan	US	社会科学	https://www.icpsr.umich.edu/web/pages/index.html	
4	Digital Curation Centre	University of Edinburgh	UK	社会科学	https://www.dcc.ac.uk/	
5	DataONE	CDLほか	US中心	地球科学	https://www.dataone.org/	
6	NIH scientific data sharing	National Institutes of Health	US	生命科学	https://sharing.nih.gov/	臨床医学に関するGood Laboratory Practice (GLP) 関連資料など
7	OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring	OECD	International	実験科学	https://www.oecd.org/en/publications/serials/oecd-series-on-principles-of-good-laboratory-practice-and-compliance-monitoring_g1gha2d5.html	
8	ESIP	NASA, NOAA and USGS	US	地球科学	https://www.esipfed.org/	
9	AGU Open Science	AGU	US	地球科学	https://www.agu.org/learn-about-agu/about-agu/open-science	地球科学における実践
10	NCEI Education	NCEI	US	環境科学	https://www.ncei.noaa.gov/resources/education	Educator向けツールキットあり
11	EarthData	NASA	US	地球科学	https://www.earthdata.nasa.gov/	分野横断的なデータ利活用の事例あり
12	KonsortSWD	NFDI4Society	GE	社会科学	https://www.konsortswd.de/en/services/research/teaching-training/	独NFDIのRDM関連ページ
13	FAIRsharing	FAIRsharing team	UK中心	生命科学	https://fairsharing.org/search?fairsharingRegistry=Standard	生命科学研究の語彙情報やガイドラインなどの標準を1800件以上収集
14	TogoTV	DBCLS	JP	生命科学	https://togotv.dbcls.jp/result.html?query=データ登録&type=manual&page=1	生命科学研究で出力されるデータのレジストリ登録などの手法を動画で共有
15	九州大学 研究データ管理支援 ページ	九州大学	JP		https://rds.dx.kyushu-u.ac.jp/research_data_management	九州大学内でのポリシーと対応手順の共有・リンク
16	お茶の水女子大学 研究支援ページ	お茶の水女子大学	JP		https://www.ocha.ac.jp/program/menu/research_support/index.html	お茶の水女子大学内でのポリシーと対応手順の共有・リンク
17	京都大学における研究データ管理 (RDM)	京都大学	JP		https://rdm.kyoto-u.ac.jp/	京都大学内でのポリシーと対応手順の共有・リンク
18	UTokyo Open Science Portal	東京大学	JP		https://sites.google.com/g.ecc.u-tokyo.ac.jp/osp/home?authuser=0	東京大学内でのポリシーと対応手順の共有・リンク。RDMkitのライフサイクル画像を引用し表示している。
19	JST RISTEX 研究データマネジメントの取組	JST	JP	社会科学	https://www.jst.go.jp/ristex/internal_research/rdm/	RDMに関する冊子データとイベントレポートへのリンク
20	名古屋大学の学術データ管理・公開・利活用	名古屋大学	JP		https://rdm.nagoya-u.ac.jp/	名古屋大学内でのポリシーと対応手順の共有・リンク
21	Galaxy Training!	Galaxy project	International	生命科学	https://training.galaxyproject.org/training-material/topics/galaxy-interface/#st-manage	生命科学データ統合解析環境Galaxyにおけるデータ管理手順
22	OMNIWeb Plus	NASA	US	地球科学	https://omniweb.gsfc.nasa.gov/	太陽圏データの取り扱い方法に関するパスファインダー。データ校正の作法を記述したテクニカルレポートへの引用など
23	RDM Compas	NFDIほか	GE	社会科学	https://rdm-compas.org/en/homepage	独社会科学分野の情報提供サイト。RDMkitにコンセプトが類似している。

欧州の「RDMkit」

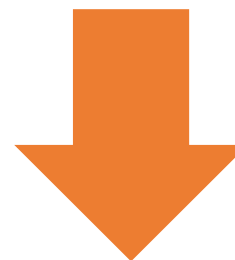
欧州 ELIXIRが提供する生命科学研究データのRDMに関するトレーニングツール



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。



<https://rdmkit.elixir-europe.org/>

国際プロジェクトELIXIR-COVERAGEでは、「RDMkit」サイトからRDMに関する情報を、CC-BYライセンスで公開している。

この枠組みを利用し、日本国内の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「RDMkit-jp」を構築

トップページ

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	識別子	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	タスク	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

<https://rdmkit.rcos.nii.ac.jp/>

体系化された内部のコンテンツ



研究データのライフサイクル

計画

- 計画（データ管理計画）とは？
- なぜデータ管理計画が重要なのか？
- データ管理計画では何を考慮すべきか？
- 詳細はこちら

計画（データ管理計画）とは？

計画（データ管理計画）とは、プロジェクト内で生成されたデータや文書を管理するために使用する計画を指します。データ管理に関する用語や手順は、プロジェクト終了後も研究データが有効な科学に貢献できるよう、事前に考えておく必要があります。

プロジェクトで生成されたデータの取り扱い、データ管理計画（DMP）で決定します。DMPは、プロジェクトの計画、途中、終了後のデータ管理プロセスについて記述した文書です。DMPの一般的な構成要素は以下の通りです。

- プロジェクトの一般的な情報
- 使用・生成されるデータセットの説明
- メタデータ、オンロジーの利用手段、データ文書の提供方法
- プロジェクト期間中および終了後のストレージソリューション、データセキュリティ、保存期間
- データの共有
- データ管理に必要なコストとリソース
- プライバシー、知的財産、ライセンスなどの権利関係、法的問題

なぜデータ管理計画が重要なのか？

研究データの「質」は研究の鍵であり、データ管理計画の立案により、自身の研究に、チームワークの促進、サービスやツールとの連携が実現します。さらに、DMPを事前に作成することで研究データをより安全なものにします。

データ管理計画を作成するメリット：

- 研究機関や資金提供者からの要求に応えられる。
- 必要なリソースや設備を計画し、経費の見積もりに役立つ。
- プロジェクトチームのデータ管理における役割と責任を明確にできる。
- データの取り扱いにおけるリスクを特定し、事前に解決策を講じることができる。
- データの活用、再利用、保存が容易になる。

データ管理計画では何を考慮すべきか？

- 研究機関や資金提供者は、プロジェクトが資金提供を受ける際に、研究資料の一部としてDMPを提出することがあります。まずは提供されるデータ管理計画のガイドライン、ポリシー、ツールなどを検討してください。
- データ管理は、研究プロジェクトの初期段階で計画する必要があります。できれば、データ生成を開始する前にDMPを作成すべきです。しかしインフラ、研究ソフトウェアの更新や新しい研究研究など、研究プロジェクトの進展に合わせて、DMPを更新する必要があります。
- 利用する施設やサービスの規約や利用条件を考慮してください。
- DMPでは多岐にわたる必要があるため、IT部門、図書館、データの管理または運用責任者、法務または情報管理チーム、データ管理担当者など、関係機関が関係する機関と連携し、支援を受けることをお勧めします。
- 本サイトで紹介している研究データ管理のベストプラクティス、ガイドライン、ツール、リソースをご覧ください。

詳細はこちら

タスク	資料
データ管理計画 How to write a Data Management Plan (DMP)	Training in TdS
データ分析 How to make data analysis FAIR	

RCOS

国立情報学研究所 オープンサイエンス推進センター
〒100-8501 東京都千代田区千代田1-8-1
www.nsl.nsl.ac.jp 03-5460-1111（受付時間：平日9:00～17:00）

About | Contact | Privacy

© 2023 National Institute of Information and Communications Technology

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

RDMkit.jp

HOME RDMkit.jpについて お問い合わせ

Google 検索

データ管理

研究データのライフサイクル

役割

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

タスク

ツール&リソース

教材

研究者

- シナリオ
- フォーカス
- はじめに
- 詳細はこちら

シナリオ

私が申請している資金提供団体は、データマネジメントプラン（DMP）を要求しています。私はDMPを書いた経験がほとんどなく、どの程度の詳細さが求められるのかわかりません。所属機関内でデータマネジメントの専門家にアクセスできる機会が限られています。データマネジメントのニーズに合わせてRDMkitの使用を検討しています。また、データ管理要件、データアーカイブ、DMPツールに関するローカルトレーニングの有用な参考資料を見つけないと思っています。

作成するデータの種類とおおよその量はわかっていますが、共同研究者とデータを共有する方法や、データを安全に保管する方法については考えていませんでした。当初は、高性能なコンピュータとクラウド環境を利用する予定でしたが、今では、国のコンピュータインフラ（例：NIIリサーチデータクラウド）を利用する必要があると考えています。私が関わっている分野では、データやキュレーションの基準が明確に定義されており、例えば、データの収集方法やサンプリング方法に関する情報（メタデータ）を取得することができます。しかし、分析に使用したツールやデータベースのバージョンなどの来歴データを保存することの重要性については、まだよくわかりません。

フォーカス

- 助成金申請のためのデータ管理計画を作成しましょう
- 法的・倫理的側面を含め、機関の方針を確実に遵守しましょう
- 適切なデータの整理と保管を行いましょう
- データの安全な共有、再現性、保存の確保をしましょう
- RDMの優れた実践例を自分のグループに伝えましょう

はじめに

- RDMライフサイクルの様々なステップ、特に計画段階を確認しましょう
- 機関のデータ管理者、またはNII RCOSの国内担当者を特定し、連絡を取りましょう
- 国または機関が提供するDMP申請ガイドラインを使用する、または入門的なトレーニングを受けましょう
- DMPを考慮してプロジェクトの計画を開始しましょう

詳細はこちら

タスク

コンプライアンスの監視および測定
Measure compliance to data management regulations and standards.

データ分析

リソースリスト

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。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

言及されている
教材のリスト

RDMkit.jp

データ管理

研究データのライフサイクル

再利用 計画 収集 加工 分析 保存 共有

役割 ▼
タスク ▼
ツール&リソース
教材

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、または間違っている場合は、「お問い合わせ」ページからお知らせください。

Show 25 entries Search: Type here...

ツールまたはリソース	記述	関連ページ
Anaconda.org	「Anaconda」は、データサイエンスと機械学習のためのオープンソースプラットフォームです。PythonやR言語のパッケージ管理とデプロイメントを容易にすることで、データ分析、可視化、機械学習のプロジェクトを効率的に実行できます。主に研究者、データサイエンティスト、開発者向けです。	データ分析
Annotare 2.0 submission tool	「Annotare」は、機能ゲノミクスデータをBioStudiesのArrayExpressコレクションに提出するためのツールです。MAGE-TAB形式を生成し、MIAME/MINSEQE/MINSCOEのガイドラインに準拠した高いメタデータ標準をサポートしています。これにより、研究者は自身の実験データを効率的に公開し、広範な研究コミュニティと共有できます。	文書化とメタデータ
Apache Maven	「Apache Maven」は、プロジェクトのビルド、レポート、ドキュメント管理を中央の情報から管理するソフトウェアプロジェクト管理ツールです。プロジェクトオブジェクトモデル (POM) のコンセプトに基づき、開発プロジェクトを効率的に管理します。	データ分析
ArrayExpress	「ArrayExpress」は、EBIによって提供される機能ゲノミクスデータのアーカイブです。このプラットフォームは、遺伝子発現実験のデータセットを集約し、研究コミュニティに公開しています。MIAME準拠の実験データとメタデータのアクセスを提供し、研究の透明性と再現性を高めることを目指しています。	文書化とメタデータ
Arvados	「Arvados」は、大規模な生物医学データの管理と処理を目的としたオープンソースプラットフォームです。データとワークフロー管理機能を統合し、ペタバイト級のデータ分析と再現可能な計算ワークフローを実行します。Arvadosは、データライフサイクル全体をサポートし、データの取得から分析、安全な共有、監査、再利用に至るまでのプロセスをカバーしています。	データ分析
Arvados	With Arvados, bioinformaticians run and scale compute-intensive workflows, developers create biomedical applications, and IT administrators manage large compute and storage resources.	保存
Bioconda	「Bioconda」は、生物情報学関連のソフトウェアパッケージをインストールするためのオープンソースのCondaチャンネルです。1400人以上のコントリビューターによって維持されており、生物学的なデータ処理や解析に特化した多様なツールを提供します。また、Condaパッケージと対応するDockerコンテナが自動的に作成され、利用可能です。	データ分析
Biocontainers	BioContainersは、バイオインフォマティクスソフトウェアのコンテナとパッケージを作成、保存、配布することを目的するオープンソースプロジェクトです。このコミュニティ主導のフレームワークは、プラットフォームに依存しない実行環境を提供し、メタデータ、バージョン、ライセンス、ソフトウェア依存関係などを標準化するためのガイドラインを開発しました。BioContainersは、バイオインフォマティクス研究のための強力な分析パイプラインを組み合わせたことが可能で、各種研究室	データ分析

RDMkit.jp

データ管理

研究データのライフサイクル

再利用 計画 収集 加工 分析 保存 共有

役割 ▼
タスク ▼
ツール&リソース
教材

教材

研究データ管理に役立つ教材のリストです。このリストは、RDMkit.jpの各ページで紹介した教材を中心に作成しています。

Show 25 entries Search: Type here...

名前	関連ページ
Data Carpentry lessons	
ELIXIR Data Management Network	
FAIR guiding principles	
GO FAIR starter kit	
Mantra - RDM training	
NPOS/ELIXIR data steward competency framework	
RDNL & DCC - Delivering RDM Services	
RDNL - Essentials for Data Support	
Science Europe - Practical Guide to the International Alignment of RDM	
TeSS - ELIXIR's training porta	
Training in TeSS	

Showing 1 to 11 of 11 entries Previous 1 Next

RCOS

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
〒101-8430 東京都千代田区一ツ橋2-1-2
rcos-ext@nii.ac.jp ※号を半角に記入してください

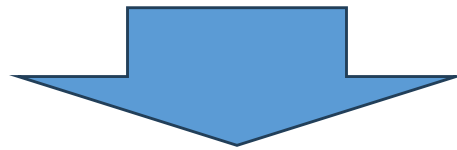
・ About
・ Contact
・ Privacy

注があるものを除いて、このサイトのコンテンツは、クリエイティブ・コモンズ 署名4.0ライセンスの下で提供されています。

© 2023 National Institute of Informatics

RDMを日本語で説明した際の**気づき**

- RDMを説明するために必要となる「語彙」が、一般や学术界でも使われないものが多く、丁寧に説明しても分かりづらい
- RDMkit-jpでは語彙についても説明しているが、冗長な記述になりがちで「直観的な理解」は難しい・・・



イラストで説明してみては？

NEW

RDM概念のイラスト化

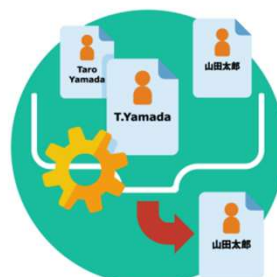
RDM用語句の概念を専門家が再検討し、
専門イラストレーターに依頼してアイコン画像を作成。
画像はサイトより配布予定



大学図書館員



教授



データ名寄せ



重複削除



データ統合



データ管理者



研究データ



DMP, 研究データ管理計画



プレプリント



大学紀要



研究成果報告書



キュレーションされたデータ

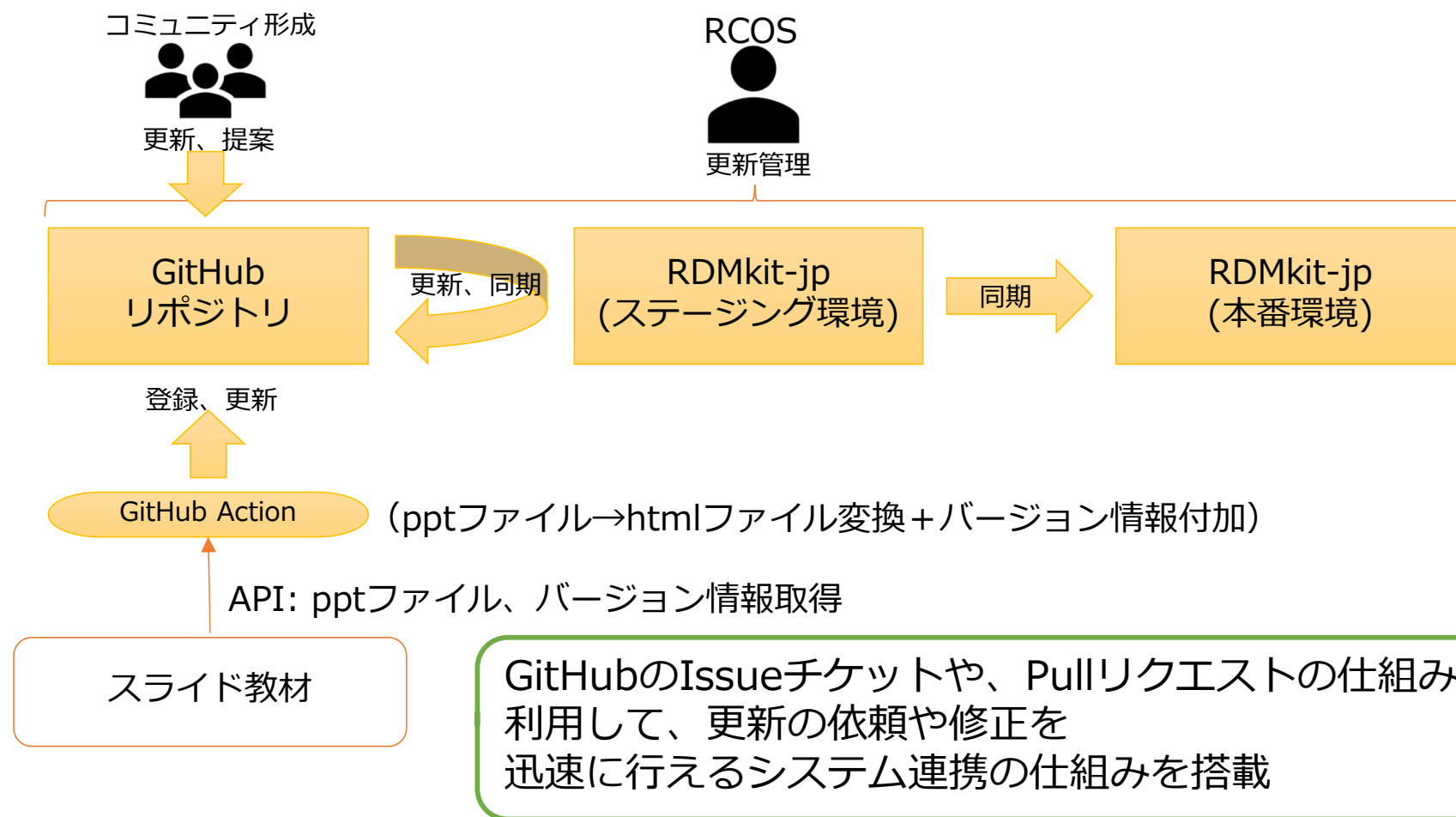


メタデータ



発表スライド

GitHubを利用した情報更新体制



より詳細なコミュニティの提案や、RDMに対する期待については、
本チームで「現場インタビュー」を実施中。
将来のコンテンツ強化に活かしていく。

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
- イラストによる、イメージしづらいRDM関連情報を追加
- 更なるコンテンツ強化で国内のRDM整備に貢献するサービスを目指す