

○大波 純一、南山 泰之、古川 雅子

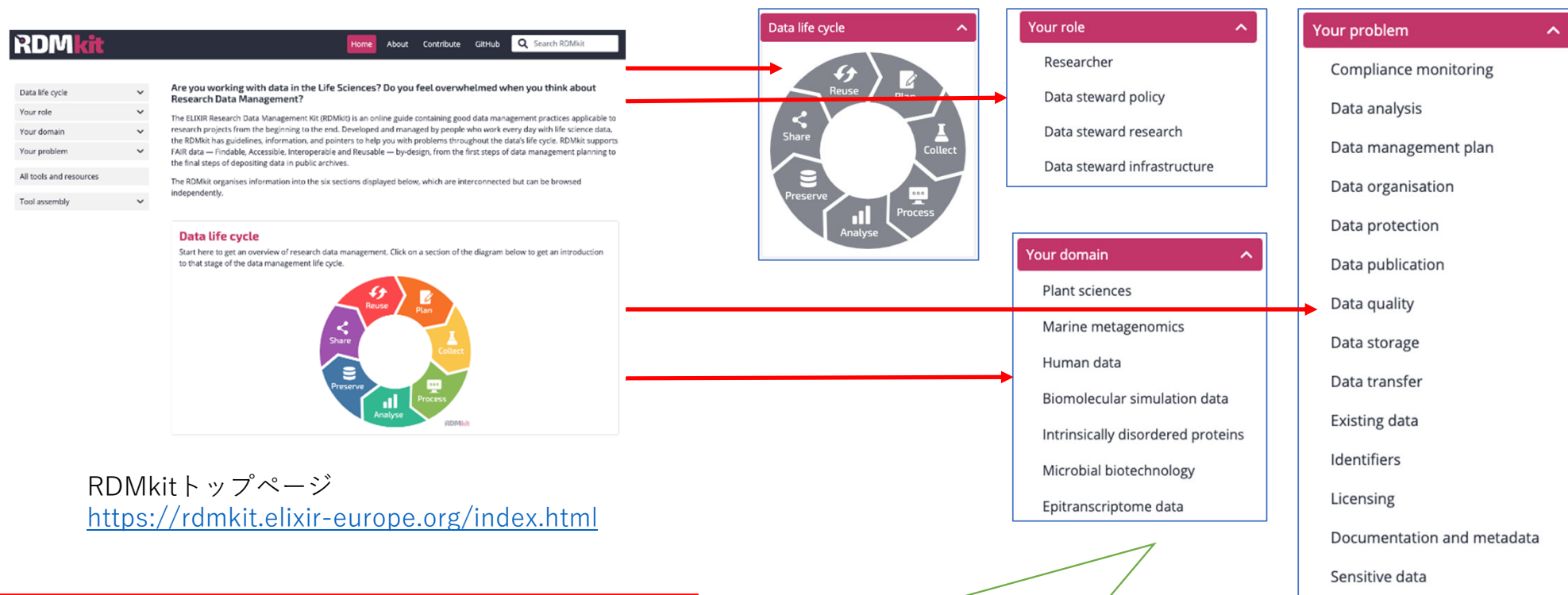
国立情報学研究所 オープンサイエンス基盤研究センター

近年、日本国内でも研究データ管理（RDM）の機運が高まりつつあり、第6期科学技術・イノベーション基本計画においても、機関ごとのデータポリシー策定や研究計画段階でのデータマネジメントプラン

（DMP）作成が急務とされている。しかしこのようなRDMに関する初心者向けの平易な表現や、ベストプラクティスの共有は国内で十分に行われているとは言い難い。一方欧州では、生命科学系リソースを共有する政府間組織ELIXIRが2021年より「RDMkit」を公開し、実務的なRDMの知識共有を実践していた。

そこでRDMに関する基盤を国内に提供している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では、RDMkitの情報を再構築し、日本語の説明や国内で利用可能な基盤を取りまとめて、新たに「RDMkit-jp」として2022年10月より公開する予定である。このサイトではRDMやDMP、識別子といった研究データ管理のために必要な語句の解説から、実際に研究データを扱う際の手順やツール情報を、関係する役割や研究データのライフサイクルごとに閲覧することができる。データを生成する研究者や研究データ管理者のみならず、学生にもアクセスし易いRDMkit-jpの情報は、日本の研究環境改善に寄与できるものと考えられる。

RDMkit: 欧州ELIXIRが提供する情報ポータルサービス



RDMkit トップページ

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

ELIXIRが提供する、バイオプロジェクト用のRDM（Research Data Management, 研究データ管理）に関するトレーニングツール

情報を「研究データライフサイクル」、「役割」、「研究分野」、「課題」に分けてフィルタして表示

RDMkitの提供する情報

Researcher Edit me

- Applies to
- Scenario
- Focus
- Getting started
- Common problems
- Relevant tools and resources

Applies to

PhD candidates, research grant applicants, project managers, group leaders, PIs

Scenario

The funding organisation I am applying to requires a data management plan (DMP). I have little experience in writing a DMP, and I am not sure of the level of detail I am required to provide. I have limited access to data management experts within my institution. I am considering using the RDMkit for my data management needs. I also hope to find useful references to local training about data management requirements, data archives and DMP tools.

I know the types and the approximate amount of data I will generate, but I have not thought about how to share data with my collaborators and how to store data securely. Initially, my plan was to buy a powerful computer and portable hard drive, but I am now thinking that I need to use a national computing infrastructure. The field I work in has well defined data and curation standards, for example, capturing information (metadata) about how to collect and sample my data. However, I am not yet familiar with the importance of storing provenance data, such as tool and database versions used in analysis.

立場や目的ごとに

- どのようなシナリオに従ってデータ管理すべきか
- 最初に手をつけるデータ管理プロセスはどこか
- よくある課題
- 関連ツール

が、自然文でまとめられている。

役割や地域ごとに異なる、RDM
(Research Data Management, 研究データ管理) の事情を理解するための需要に応える内容

様々な国や学術機関が提供するRDM情報

2021年10月に”Research Data Management”でGoogle検索しヒットした情報を30件目視し、教育のための解説サイトを抽出した。

ページ名 / 組織名	地域	URL
Open University, Library research support	イギリス	https://www.open.ac.uk/library-research-support/research-data-management
これからのリサーチデータマネジメント (RDM) / 京都大学	日本	https://www.kulib.kyoto-u.ac.jp/uploads/RDM_leaflet_201911.pdf
研究データマネジメントとは / 京都大学	日本	https://www.kulib.kyoto-u.ac.jp/researchdata/1380894
Sheffield university library	イギリス	https://www.sheffield.ac.uk/library/rdm
University of Pretoria	南アフリカ	https://up-za.libguides.com/c.php?g=356288
Maastricht University	オランダ	https://library.maastrichtuniversity.nl/research-support/rdm/
OCLC	国際組織	https://www.oclc.org/research/areas/research-collections/rdm.html
Ghent University	ベルギー	https://www.ugent.be/en/research/datamanagement
Ottawa大学	カナダ	https://biblio.uottawa.ca/en/services/faculty/research-data-management
University of Pittsburgh	アメリカ	https://pitt.libguides.com/managedata

国や組織ごとにRDM情報をまとめ、提供する動きが盛んになりつつある。

RDM情報提供の重要性

- RDMの知識やノウハウを、日本語の平易な言葉で説明する専門のサイトは日本にはまだ存在しない。
- 欧州RDMkitはサイトの構成をCC-BYで無償提供している。
- 今後、研究データの管理や知識はさらに重要性を増すことが予想される。



- 広範な日本の研究データ管理を実施している、
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で
RDM情報を収集
- RDMkitをベースとしたポータルサイト「RDMkit-jp」から公開

RDMkit-jp サイト

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

The screenshot shows the RDMkit-jp website interface. At the top, there is a navigation bar with links for HOME, 概要 (Overview), 貢献する (Contribute), お問い合わせ (Contact), and GitHub. A search bar is also present. On the left, a sidebar menu titled 'データ管理' (Data Management) lists various sections: 研究データのライフサイクル (Research Data Lifecycle), 役割 (Roles), タスク (Tasks), 研究データ管理基盤 (Research Data Management Platform), ツール&リソース (Tools & Resources), and 教材 (Materials). The main content area is titled '研究データ管理、行っていますか?' (Are you managing research data?). It contains a paragraph explaining RDMkit-jp's purpose and a circular diagram illustrating the '研究データのライフサイクル' (Research Data Lifecycle). The lifecycle is represented by a circle divided into six colored segments, each with an icon and a label: 再利用 (Reuse) in red, 計画 (Plan) in orange, 収集 (Collect) in yellow, 加工 (Process) in green, 分析 (Analyze) in blue, and 共有 (Share) in purple. Below the diagram, there is a text box explaining that clicking on the segments will lead to detailed explanations of each stage.

ソースはGitHubで管理し、課題も受ける。

トップページに、「研究データ管理」に関する、平易な説明と案内。

左側に折りたたみ式的全項目のメニューを設置。

図形の中から直接クリックして、各フェーズの説明へ移動

RDMkit-jpサイトのカテゴリメニュー

役割
研究データ管理における自分の役割を確認し、自分に関連する研究データ管理リソースを見つけ、キャリアパスの進展に役立つ情報を得ることができます。
[Show pages ▼](#)

タスク
研究データ管理に関する一般的な問題に取り組むためのガイドラインと解決策を見つけましょう。
[Show pages ▼](#)

研究データ管理基盤
研究データ管理を統合的に実践するためのソフトウェアを紹介しています。

- [研究データ管理基盤](#)

ツール&リソース
研究データ管理に役立つツールやリソースを集めたカタログを紹介しています。

- [ツール&リソース](#)

教材
研究データ管理に関する教材のカタログを紹介しています。

- [教材](#)

新着情報
Linking with DSW
📅 2022年2月28日 10:00 🗣️ 803
This page gives an insight in RDM on national level in France. Discover the page here [こちらのページをご覧ください](#)。
その他のニュースについては [ニュースページ](#)をご覧ください。

各大大項目の説明とリンク

RDMkit-jpサイトの統計情報



本家RDMkitに倣い、
CC-BYとする。

説明ページ内の章立て



計画・収集・加工・
分析・保存・共有・
再利用の各フェーズ
ごとに説明

研究データのライフサイクル プランニング

- データマネジメント・プランニングとは？
- 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・プランニングでは何を考慮すべきか？
- 詳細はこちら

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、プロジェクトで発生したデータやドキュメントの管理に使用する予定の戦略を定義することです。これは、データ管理に関する懸念や予想外のコストを回避するための最良の方法を事前に考え、プロジェクト終了後研究データが科学の分野で可能な限り大きな価値を発揮できるように条件を整えることです。

プロジェクトで生成されたデータの取り扱いに関するソリューションは、通常、データマネジメントプラン (DMP) で公式化されます。DMPは、プロジェクトの前、中、後に発生するデータ管理プロセスのいくつかの側面を記述した文書です。DMPが一般的な構成要素は以下の通りです。

- プロジェクトに関する一般的な情報
- 使用・生成されるデータセットの説明
- メタデータ、イントロダクション、データ文書の提供方法
- プロジェクト期間中は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、保存戦略
- 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、ライセンスなどの倫理的・法的懸念

なぜデータマネジメント・プランニングが重要なのか？

研究データを大切に、データ管理計画を立てることは、良い研究習慣です。これにより、作業の効率化、チームワークの促進、サービスやツールの利用が可能になります。さらに、詳細なDMPは、研究データをより公正なものにするのに役立ちます。データマネジメントプランを作成する理由は：

- 研究機関や資金提供者からの要求であることが多い。
- 必要なリソースや設備の計画と予算化に役立つ。
- 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する。
- データの取り扱いにおけるリスクを特定し、早期に解決策を講じることができる。
- データの共有、再利用、保存が容易になる。

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？

- 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、助成金申請書の一部として、または最終のものとしてDMPを要求することがよくあります。したがって、資金提供者が要求するデータマネジメント計画のガイドライン、ポリシー、ツールを検討してください。
- データマネジメントは、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。望ましくは、データ収集を開始する前にDMPを取り入れるべきです。しかし、DMPは生きた文書であるため、研究プロジェクトの進行に合わせて、インフラ、研究ソフトウェアの更新や新しい研究などに合わせて更新する必要があります。
- 使用を予定している施設やインフラで必要とされる基準やベストプラクティスを考慮してください。
- DMPでは様々な側面に対応する必要があります。IT部門、図書館、データマネージャーまたはデータスチュワード、法務または技術サポートチーム、データ保護担当者など、所属機関のサポートサービスから推奨事項を見つけ、支援を受けるのが良いです。
- 本サイトでは、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、リソースを紹介しています。

詳細はこちら

タスク	資料
データマネジメントプラン How to write a Data Management Plan (DMP)	Training in TeSS
データ分析 How to make data analysis FAIR	

平易な日本語でのRDMに関する説明。

理由付けや考え方を、利用者の立場に立って説明する。

関連する項目や、説明の中で言及があったツールを下部にリンク。

役割ごとのページ

RDMkit.jp

HOME 概要 貢献する お問い合わせ GitHub

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

A FAIR guide for data providers to maximise sharing of human genomic data

タスク

研究データ管理基盤

ツール&リソース

教材

役割

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盤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元データ保存することの重要性については、まだよくわかっていません。

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.

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

の4ランクに分け、シナリオやToDoを説明。

タスクごとのページ

タスク
データマネジメントプラン

- データマネジメントプラン (DMP) を起草するために、どのようなテンプレートを使うべきか？
- DMPを書くためにどのようなツールを使うべきか？
- DMPには何を書くべきか？
- 詳細はこちら

データマネジメントプラン (DMP) を起草するために、どのようなテンプレートを使うべきか？

概要

現在、様々な資金提供者や機関から、数多くのDMPテンプレートが提供されています。さらに、machine-actionable DMPのテンプレートを開発する努力も続けられています。

考察

- 各資金提供機関は、特定のDMPテンプレートを要求または推奨することができます。
- あなたの所属機関がDMPテンプレートを要求・推奨することも可能です。
- テンプレートは、テキスト形式の質問リストとして、またはmachine-actionableな形式で提示することができます。

ソリューション

- DMPテンプレートが必要としているか、または推奨しているかどうかについては、資金提供機関や所属機関の文書を参照するか、またはそれらに問い合わせてください。
- 近年、日本の研究助成機関で提供されている DMPテンプレートは以下になります。
- (日本) 科学技術振興機構「戦略的創造研究推進事業におけるデータマネジメント実施方針」(2016)
- (日本) 日本医療研究開発機構「データマネジメントプランの義務化について」(2020)
- このH2020オンラインマニュアルでは、DMPガイドラインを読むことができ、ページ下方では、様々なフォーマットのH2020 DMPテンプレートをダウンロードすることができます。
- machine-actionable DMPテンプレートを作成したい場合は、Research Data AllianceのDMP Common Standardモデルの採用を検討しましょう。

DMPを書くためにどのようなツールを使うべきか？

概要

DMPは、ダウンロードしたテンプレートをテキスト文書形式で使用することで、オフラインで書くことができます。しかし、現在、ウェブベースのDMPツールが多数利用可能で、通常、いくつかのDMPテンプレートが含まれており、質問の解釈や回答のガイダンスが提供されているため、このプロセスが非常に容易になっています。また、いくつかのツールでは、DMPを共同で作成することができ、DMPが生きた文書であることから、その進捗状況を確認することができます。

識別子、既存データ、文書化とメタデータ、ライセンス、データ管理、データ品質、データ分析、データ公開、データマネジメントプラン、データストレージ、データの仲介、センシティブデータ、コンプライアンスの監視および測定、の項目に分けて説明。

欧州版RDMkitの考え方を踏襲しつつ、日本国内のツールや考え方に置き換えて表現

リソースリスト



文の中で言及されているツールや、有用と考えられるリソースをリスト化し分類。リンク集ポータルとしても利用できるように。

RDMkit-jpのまとめ

RDMkit日本版のポイント：

- これまで項目によっては論文を参照しないといけなかった内容にWebからアクセスでき、初学者の教育にも利用可能。
- 国内の各学術基盤への道案内やサポート。
- 本家RDMkitは欧州とバイオ分野に特化した内容だったが、チームで書き換えローカライズ。

今後：

- 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。
- コミュニティとの連携を検討中。本家RDMkitだけでなく、国内のRDM情報を必要とする分野と協調体制を確立していく。