

GakuNin RDM上のDMPメタデータ機能と 共有プロジェクト型DMR、 将来ビジョン

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
2023年5月30日

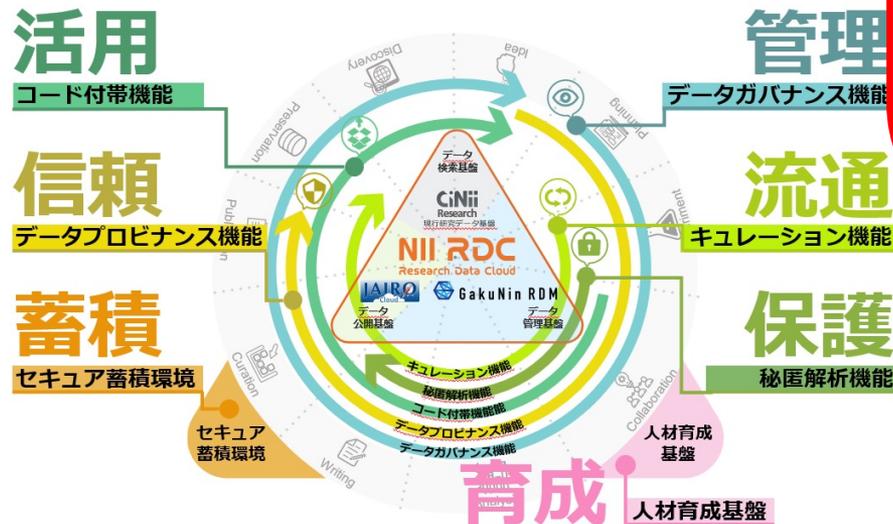
込山悠介

NII RDCでのDMRの実現を目指して

これまでの研究データ基盤 NII RDC



これからの研究データ基盤 NII RDC



提供中

現行の研究データ管理基盤
GakuNin RDM

提供中

公的研究費のデータ管理
利活用のメタデータ

第1世代DMP

機能評価試験中

データガバナンス機能

第2世代DMP

学術機関の研究データポリシーや
DMP/DMRに対応する
データガバナンスの実現を提案

GakuNin RDMと関係するDMP/DMR

- **第0世代DMP（それ以前）**
 - 公的研究費の申請書や報告書で資金配分機関が、研究データの情報を把握するために利用。ExcelやWordで作成可能なデータ管理情報の書類。データカタログなどで公開。
 - 省庁や資金配分機関毎にフォーマットがバラバラ。
- **第1世代DMP**
 - 現行のGakuNin RDMのメタデータ登録機能
 - 政府やFA主導のDMP（公的資金の研究データ利活用・公開のための共通メタデータ項目）
 - 研究データを管理しながらメタデータ管理
- **第2世代DMP**
 - GakuNin RDMにおけるデータガバナンス機能
 - DMPだけでなく、データガバナンス全般に渡りデータ管理の品質を向上させる
 - 研究プロセスの起点から終点まで、NII RDCの研究データ基盤全体に渡り、システム連携可能(machine actionable)化によるモニタリング

第0世代DMPの例：NEDOデータカタログ

NEDOデータカタログ

2022年8月

公表可能データ										
データNo.	提出日	データ名称	データの説明	管理費	プロジェクト終了後のリポジトリ	概略データ量	研究データの利活用・提供方針	連絡先	契約件名（プロジェクト名）	備考
124	2022/7/12	高分子論文集のテキスト化データ	高分子論文集のpdfファイルからテキスト部分を抽出し、文章構造についてタグ付けを行なったxml形式のデータ（外注分も含む）	国立研究開発法人理化学研究所 革新知能統合研究センター	左記に保存	1 GB	研究目的に限り一般に提供可とする。	国立研究開発法人理化学研究所 革新知能統合研究センター 知識獲得チーム チームリーダー 松本 裕治 (D.Eng.) e-mail: yuji.matsumoto@riken.jp	超先端材料超高速開発基盤技術プロジェクト/計算機支援次世代ナノ構造設計基盤技術/材料データ例構造化A Iツール開発	
123	2022/7/12	今分子論文集のアノテーション用データ	高分子論文集からテキスト抽出後にアノテーションツールの入力形式（tsvファイル）に変換したデータ	国立研究開発法人理化学研究所 革新知能統合研究センター	左記に保存	1 GB	研究目的に限り一般に提供可とする。	国立研究開発法人理化学研究所 革新知能統合研究センター 知識獲得チーム チームリーダー 松本 裕治 (D.Eng.) e-mail: yuji.matsumoto@riken.jp	超先端材料超高速開発基盤技術プロジェクト/計算機支援次世代ナノ構造設計基盤技術/材料データ例構造化A Iツール開発	
122	2022/7/12	高分子論文集から作成した教師データ・コーパス	-課題2に教師データ用として提供するためにアノテーションを実施したtsv形式のデータ（外注分も含む） -課題2で開発したAIツールの動作検証実施用として数部分を公開する	国立研究開発法人理化学研究所 革新知能統合研究センター	左記に保存	1 GB	研究目的に限り一般に提供可とする。	国立研究開発法人理化学研究所 革新知能統合研究センター 知識獲得チーム チームリーダー 松本 裕治 (D.Eng.) e-mail: yuji.matsumoto@riken.jp	超先端材料超高速開発基盤技術プロジェクト/計算機支援次世代ナノ構造設計基盤技術/材料データ例構造化A Iツール開発	
121	2022/7/12	教師データ・コーパス	-課題2に教師データ用として提供するためにアノテーションを実施したtsv形式のデータ（外注分も含む） -上記以外のもの	国立研究開発法人理化学研究所 革新知能統合研究センター	左記に保存	1 GB	研究目的に限り非営利団体に所属する個人に提供可とする。営利団体に所属する個人については、一定期間後に提供可とする。	国立研究開発法人理化学研究所 革新知能統合研究センター 知識獲得チーム チームリーダー 松本 裕治 (D.Eng.) e-mail: yuji.matsumoto@riken.jp	超先端材料超高速開発基盤技術プロジェクト/計算機支援次世代ナノ構造設計基盤技術/材料データ例構造化A Iツール開発	
120	2022/6/22	新概念モジュールに対して実施した加速試験によって得られる太陽電池特性データ	新概念モジュールに対して実施した加速試験によって得られる太陽電池特性データ（電流、電圧、曲線因子、出力）、エレクトロニクス像のデータ、加速試験条件のデータ	国立大学法人新潟大学	学内に保存	1.0 GB	提供の可否および条件は個別に相談	増田 淳 (新潟大学) a-masuda@eng.niigata-u.ac.jp	N E D O 先端研究プログラム/エネルギー・環境新技術先導研究プログラム/新概念結晶シリコン太陽電池モジュールの開発	
119	2022/6/16	次世代配線候補材料の電気抵抗に関するデータ	候補材料の膜厚および配線幅に依存した電気抵抗率変化に関するデータ	国立大学法人東北大学	担当研究室に保存	1G	提供の可否および条件は個別に相談	東北大学大学院工学研究科 知能デバイス材料学専攻 小池研究室 022-752-2299 koikej@material.tohoku.ac.jp	N E D O 先端研究プログラム/新産業創出新技術先導研究プログラム/ポストムーア時代の次世代配線開発	
118	2022/6/16	次世代配線候補材料の密着強度、拡散/バリアに関するデータ	候補材料の絶縁層に対する密着強度、および高温・高電界条件下における電気特性の変化に関するデータ	国立大学法人東北大学	担当研究室に保存	1G	提供の可否および条件は個別に相談	東北大学大学院工学研究科 知能デバイス材料学専攻 小池研究室 022-752-2299 koikej@material.tohoku.ac.jp	N E D O 先端研究プログラム/新産業創出新技術先導研究プログラム/ポストムーア時代の次世代配線開発	

https://www.nedo.go.jp/activities/tsc_data_catalog.html

第1世代DMP

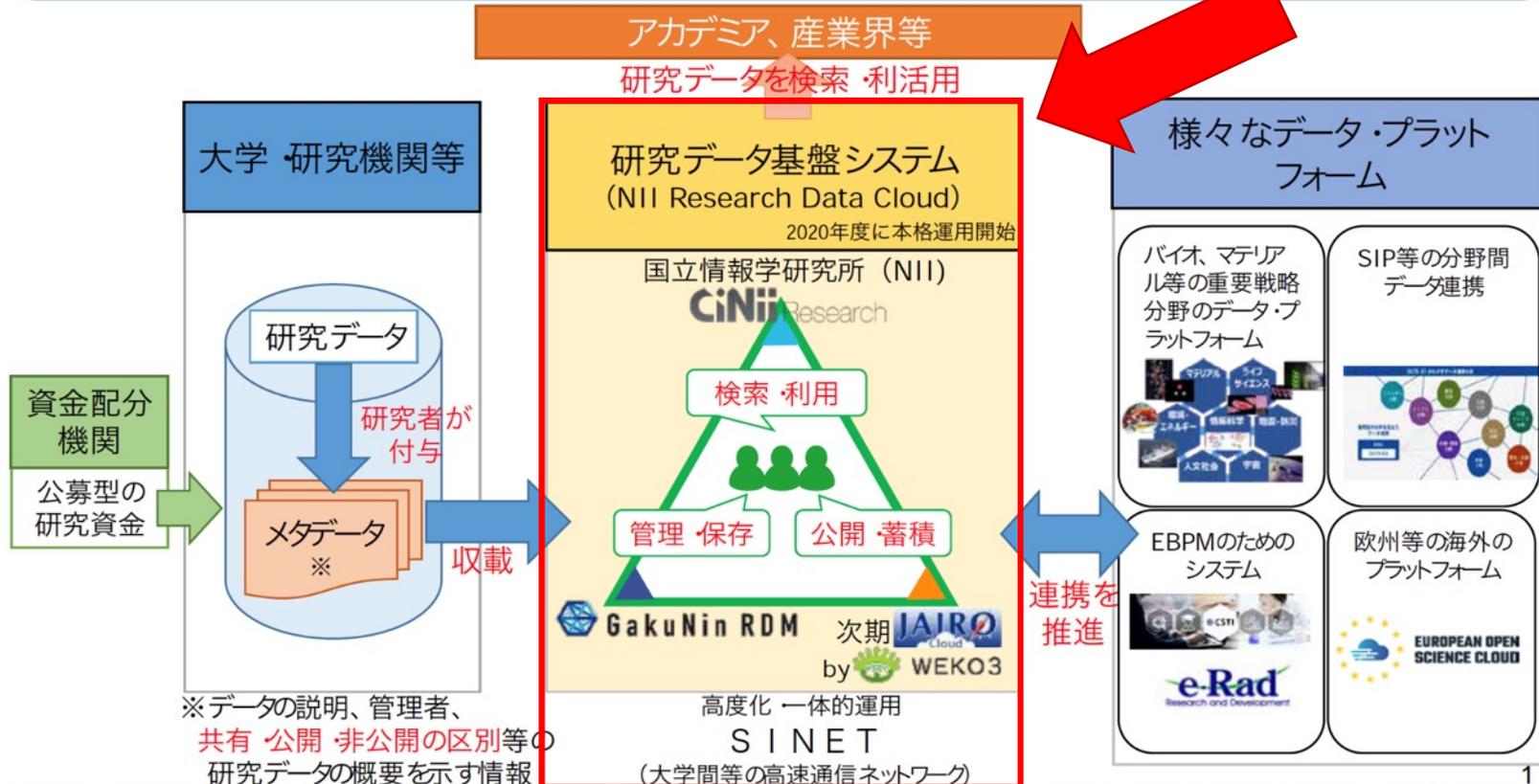
現行GakuNin RDMの メタデータ登録機能

公的資金による研究データ管理・利活用に関する基本的な考え方について (2021年)

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ついて

研究データ基盤システムを中核としたデータ・プラットフォームの構築

- 研究データの公開・共有を推進、産学官のユーザが**データを検索可能**
 - ムーンショット型研究開発制度**における試行(2020年度開始)、その後、次期**SIP**に導入
- **全ての公募型の研究資金**の新規公募分に導入(2023年度まで)



「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

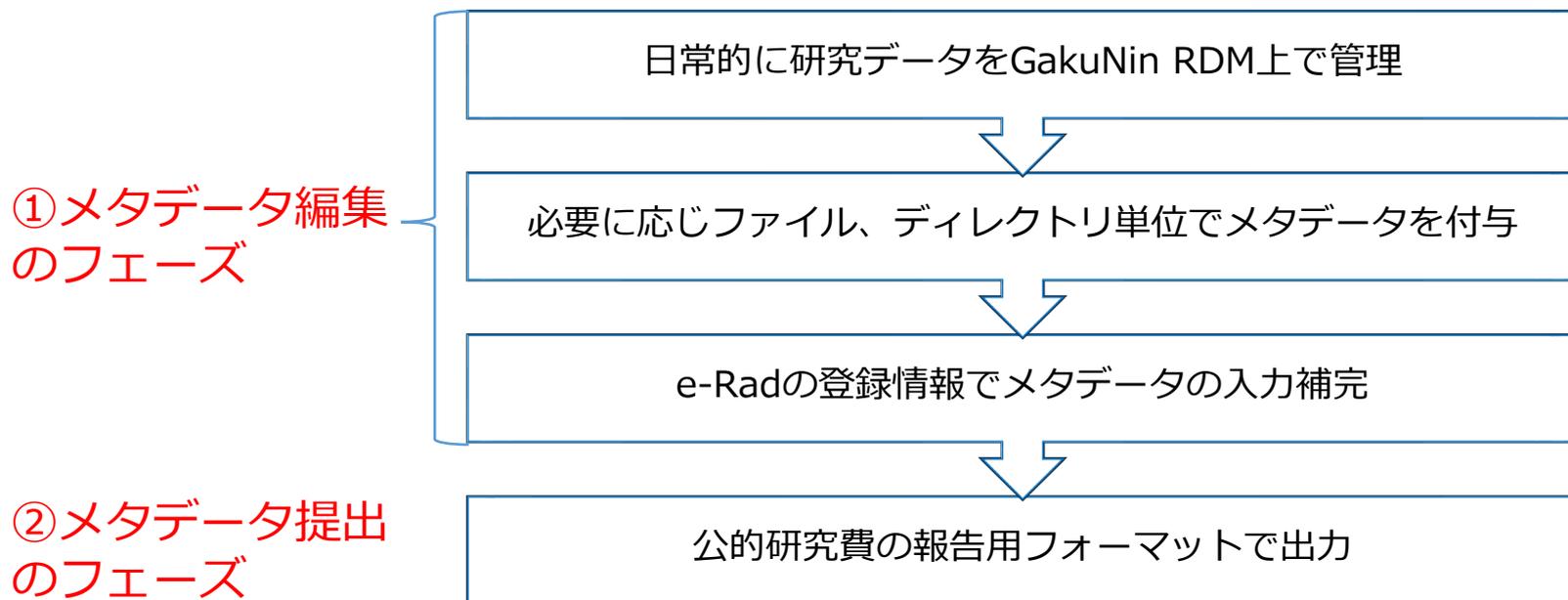
(2023年3月31日時点)

	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）の英語略称 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
	プログラム名	任意	競争的研究費制度の名称
3	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な番号 公募型の研究資金以外の場合は記入不要
	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	管理対象データを一意に特定するための番号 公募型の研究資金を配分した資金配分機関が付与 公募型の研究資金以外の場合はデータ管理機関（項目14）が付与
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	メタデータの掲載日・掲載更新日
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）。e-Radとの連携により、自動入力される予定。
9	データ種別	必須	研究データ基盤システム上では、通常は「データセット」を標準とするが、データの特性に応じて「データセット」以外の種別を選択可能
10	概略データ量	任意	1GB未満、1GB以上10GB未満、10GB以上100GB未満、100GB以上等の区分により記載
11	管理対象データの利活用・提供方針	必須	無償/有償、ライセンス情報、その他条件（引用の仕方等）等を記載
	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
12	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
	リポジトリURL・DOIリンク	任意	情報があれば記載
13	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
	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
14	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
	データ管理者	必須	データ管理組織において各管理対象データを管理する担当者の名前
	データ管理者のe-Rad研究者番号	任意	管理者のe-Radの研究者番号 e-Rad研究者番号がない管理者は記入不要、ある場合は番号研究者が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

「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」のメタデータへのデータ基盤の対応

- 基本的考え方におけるメタデータの共通項目は、公募型研究資金による研究データに付与するメタデータ項目 (15項目)を一般的に定めたもの。
- ムーンショット(MS)型研究開発制度のメタデータ項目を原案に一般化されたメタデータ。
- 先行してMS型研究開発制度におけるデータガバナンスのために使用されている。
- 公募型研究資金による研究成果のデータの管理・利活用への利用の他、内閣府や資金配分機関への報告や成果の分析等にも活用される。



GakuNin RDM中での公的資金による 研究データのメタデータ管理機能 2

①メタデータ編集
のフェーズ

ファイルメタデータの編集

メタデータ様式: 公的資金による研究データのメタデータ登録

クリップボードから貼り付け

データ No.
12345

データの名称 (日本語)*
ムーンショット型研究開発制度・サンプルデータ

Title (English)*
Sample Data of Moonshot Project

掲載日・掲載更新日*
サンプル表示
2023-05-31

データの説明 (日本語)*
サンプル表示
ムーンショット型研究開発制度のサンプルデータ

「公的資金による研究データ管理・利活用のためのメタデータ項目」
のデータスキーマへ対応。メタデータガイドライン第三版に対応

GakuNin RDM中での公的資金による 研究データのメタデータ管理機能 4

②メタデータ提出
のフェーズ

GakuNin RDM

検索 Yusuke Komiyama

登録

戻る

自動保存済み:
21 minutes ago

メタデータ登録

資金配分機関情報

JST

体系的番号におけるプログラム情報コード

12345

プログラム名 (日本語)

ムーンショット型研究開発制度

Program name (English)

Moonshot Project

体系的番号

12345678

プロジェクト名 (日本語)

サンプルプロジェクト

Project name (English)

Sample Project

プロジェクトの分野

189

登録データ

登録データ一覧

タイトル	パス	管理者
ムーンショット型研究開発制度・サンプルデータ	osfstorage/実験データサンプル/	-

管理対象のファイルやディレクトリをメタデータ管理機能で登録した際の画面イメージ。

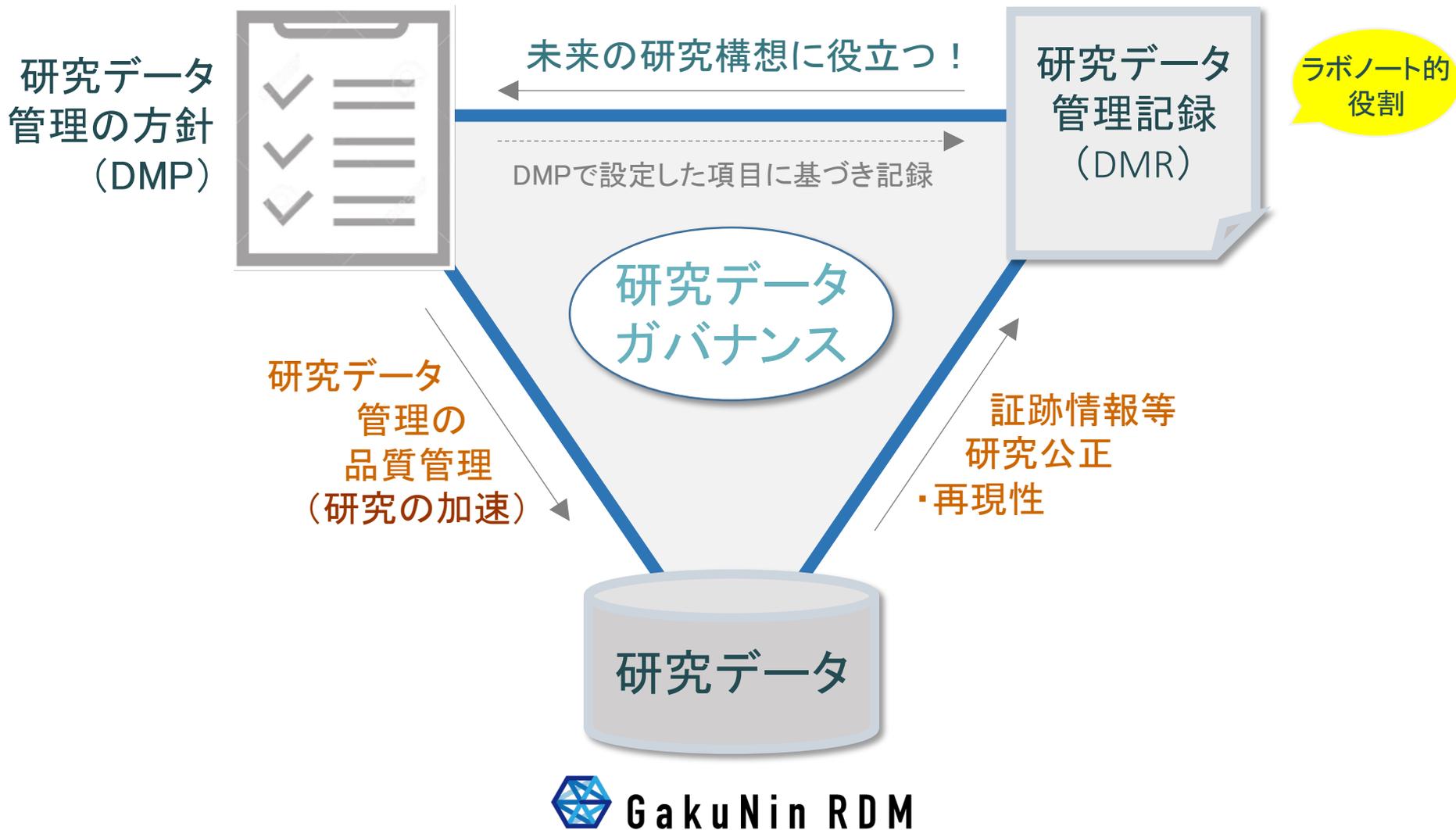
大型研究プロジェクトでの利用拡大 Moonshot型研究開発制度・目標2合原プロジェクト

The screenshot displays the GakuNin RDM web interface. The main page shows a list of components (コンポーネント) for a project, including shared domains for various projects like '合原PJ共有領域' (425 members), '片桐PJ共有領域' (424 members), '高橋PJ共有領域' (425 members), '松浦PJ共有領域' (426 members), and '大野PJ共有領域' (425 members). A modal window titled 'ファイルメタデータの編集' (Edit File Metadata) is open, showing fields for 'データ No.' (12345), 'データの名称 (日本語)*' (Moonshot型研究開発制度・サンプルデータ), 'Title (English)*' (Sample Data of Moonshot Project), '掲載日・掲載更新日*' (2023-05-31), and 'データの説明 (日本語)*' (Moonshot型研究開発制度のサンプルデータ).

作業部会を設置して、大規模（425名以上）な研究プロジェクト参加者のデータ管理・共有、メタデータ登録。複数機関・複数研究課題にまたがる研究データ管理。

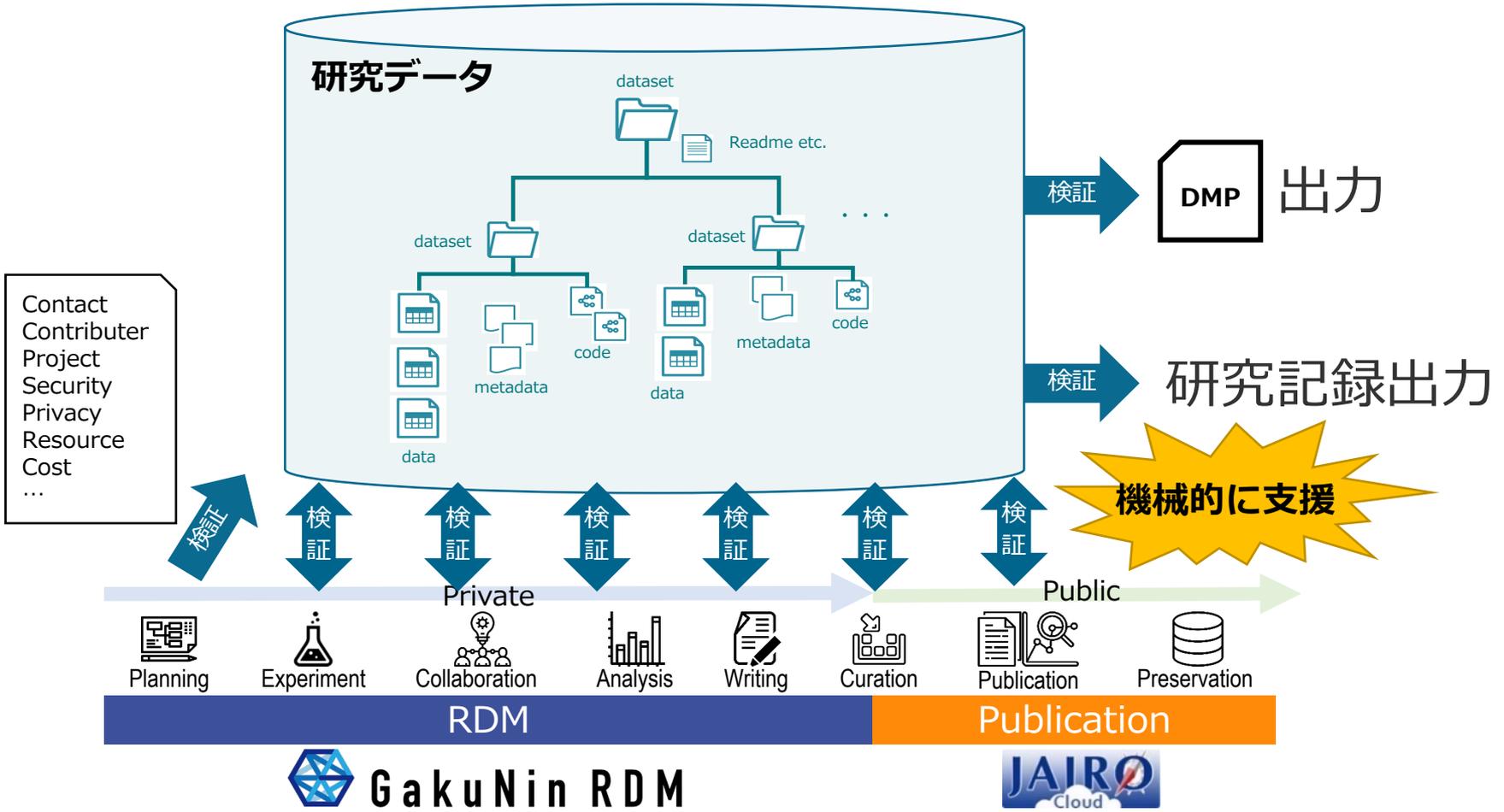
第2世代DMP GakuNin RDMの データガバナンス機能

機関における研究データの管理にむけて GakuNin RDMを用いたDMRの提案



GakuNin RDMのデータガバナンス機能

研究データの状態（＝メタデータ・データセット）の変化をDMPに基づき検証、検証プロセスを保存・提供



データガバナンス機能でのモニタリング

研究者

DMPで管理側とコミュニケーション
(データガバナンス機能を使って
DMPに沿ったデータ管理を実施)

研究管理者

FA

DMPで研究者とコミュニケーション
(研究データ管理の実態との整合性は
データガバナンス機能が保証していると期待)

DMP

データガバナンス機能

生成

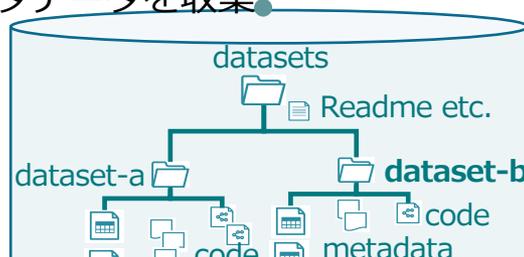
プランの実行に必要な制約

検証

メタデータのコレクション(状態)

システムから
メタデータを収集

研究者



検証ルールをスキーマで表現。

分野ごとの検証ルールのユニーク性は、それぞれのコミュニティでスキーマを拡張する形で表現できる。

DMPで定義した状態をシステム中で満足しているか検証

NII RDCを用いた学術機関のデータポリシーに基づくデータガバナンスの実現

研究者

- 研究データとその管理方針・記録(DMP/R)を機関と共有
- 研究成果と研究データを可能な限り公開

機関

- 研究データの機関管理とコンプラ対応
- 研究者の研究データの管理・公開のための環境整備

