

公的資金による研究データの 管理・利活用に関する進捗と事例 ～研究データ2024～



令和7年2月

内閣府 科学技術・イノベーション推進事務局

目次

1. 公的資金による研究データの管理・利活用の推進	P. 3
(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（概要）	P. 4
(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」に示されたマイルストーン	P. 5
(3) 公的資金による研究データマネジメントの実現のための3つのアプローチ	P. 6
(4) 研究者が取り組むこと	P. 7
(5) 研究開発を行う機関で取り組むこと	P.10
(6) 資金配分機関で取り組むこと	P.11
2. 実施状況（統合イノベーション戦略2024）	P.13
3. メタデータ説明書について	P.17
4. ムーンショット型研究開発制度における取組事例	P.18
5. 関係府省における取組事例	P.29

1. 公的資金による研究データの管理・利活用の推進

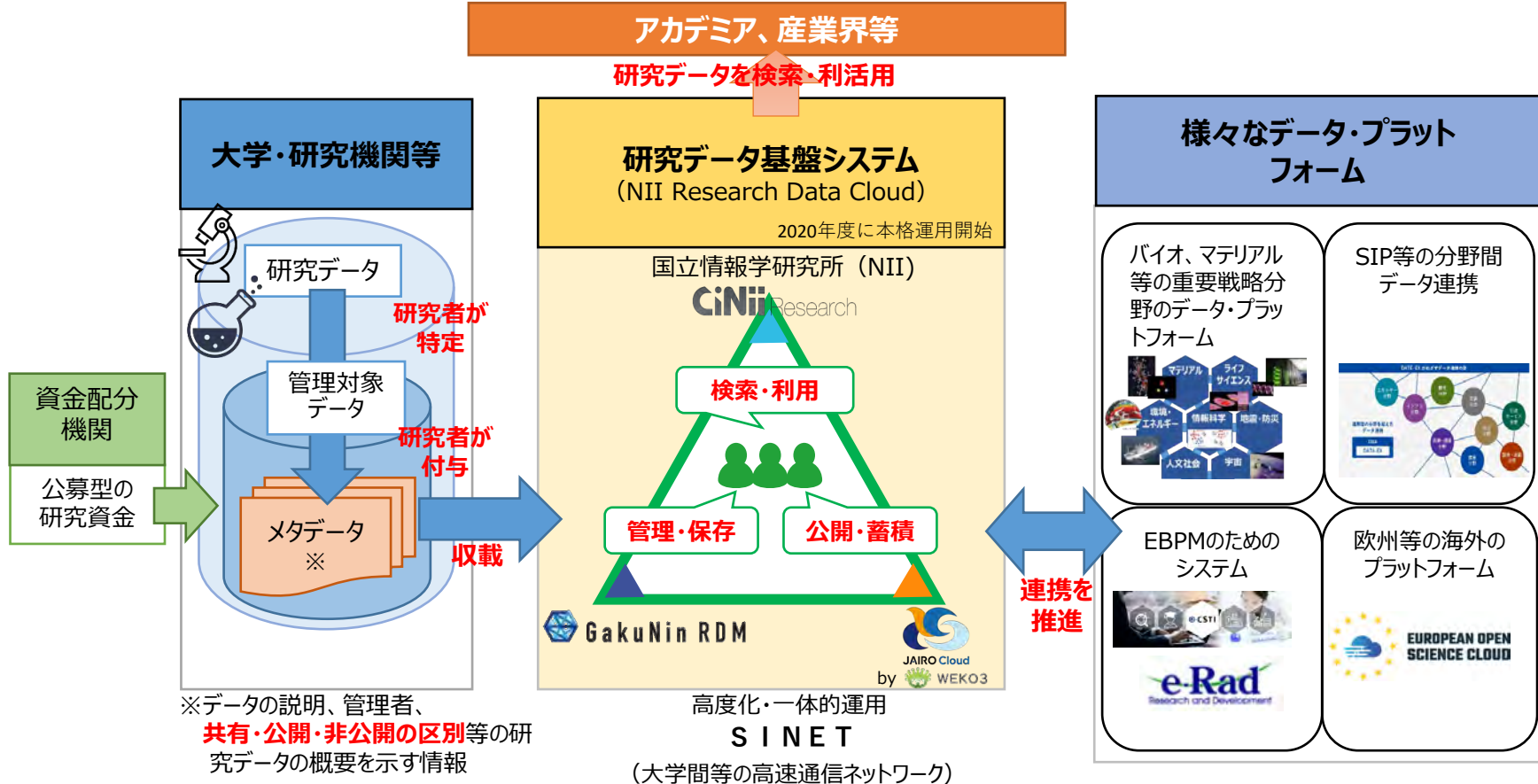
【背景】

- 知識をオープンにし、研究の加速化や新たな知識の創造などを促す**オープンサイエンスの動き**が活発化
- 世界的な出版社やIT企業**が、研究成果や研究データを**ビジネスの対象**として焦点を当てつつある

【政策文書】

- 第6期科学技術・イノベーション基本計画（2021年3月）
- 統合イノベーション戦略2024（2024年6月）
- 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（2021年4月）

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**



1.(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（概要）

<https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/sanko1.pdf>

<主な内容>

- **研究データ基盤システム（NII Research Data Cloud）**を中核的なプラットフォームとして位置付け、産学官における幅広い利活用を図るため、メタデータ（データを説明するための情報から構成されるデータ）を検索可能な体制を構築する。（2023年度まで）
- **研究開発を行う機関**は、データポリシーを策定し、機関リポジトリへの研究データの収載を進める。※1
- **公募型の研究資金**の全ての新規公募分について、メタデータを付与する仕組みを導入。（2023年度まで）※2
- **研究者**は、所属機関のデータポリシーや公募型の研究資金における資金配分機関の基準等に基づき、管理対象データの範囲を定め、メタデータを付与し、研究データ基盤システム上において検索可能となるように登録する。
- その他（人材・支援体制の整備、取組状況の評価、他のデータ・プラットフォームとの連携等）

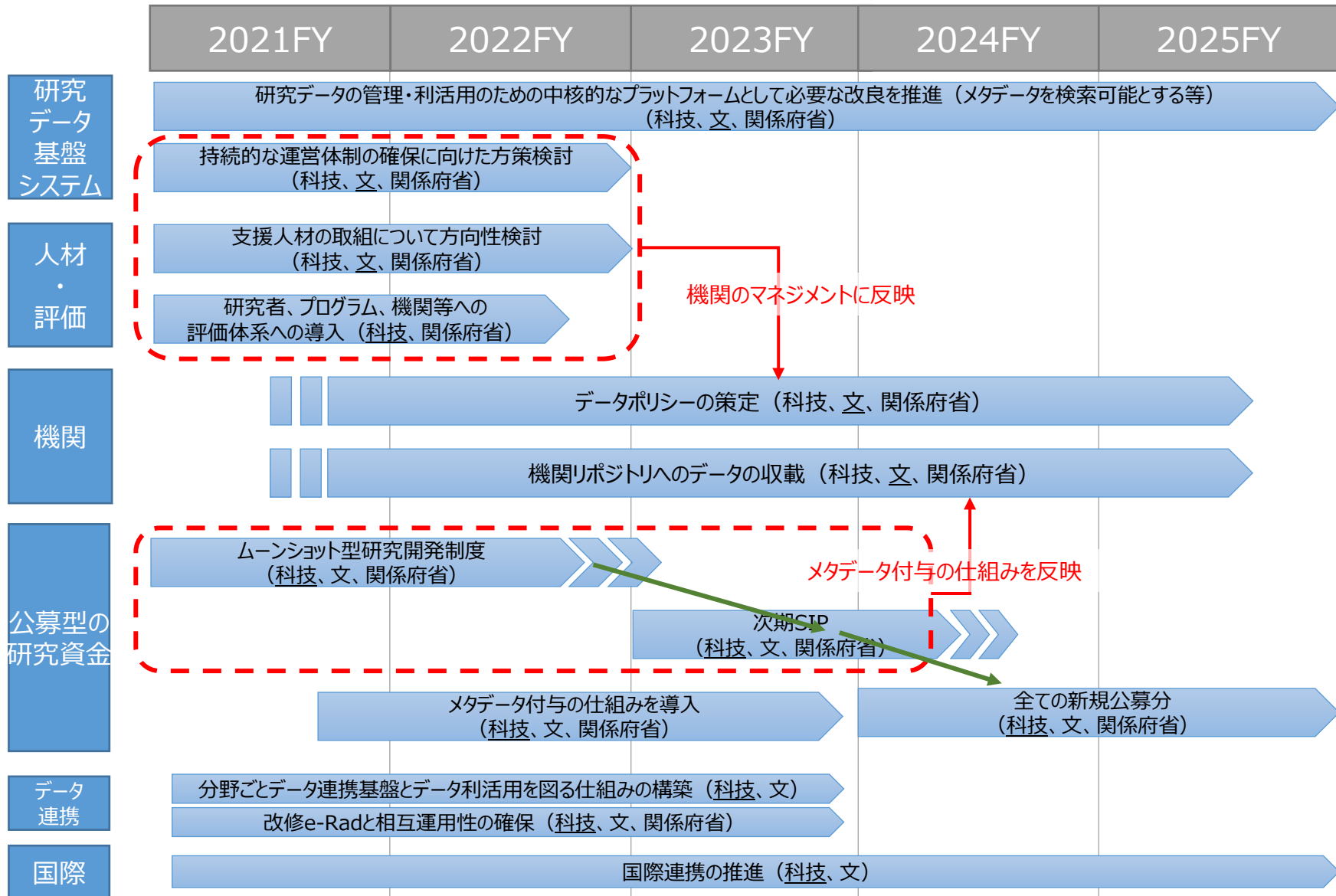
※1：国立大学法人、大学共同利用機関法人及び国立研究開発法人は、基本的な考え方において、2025年までにデータポリシーを策定することとされている。なお、国立研究開発法人については、資金配分機関である日本医療研究開発機構（AMED）、科学技術振興機構（JST）及び新エネルギー・産業技術総合開発機構（NEDO）を除く。

※2：関係府省間の合意により、競争的研究費を対象としている。

<メタデータの共通項目 最新版>

https://www8.cao.go.jp/cstp/common_metadata_elements.pdf

1.(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」に示されたマイルストーン



1.(3) 公的資金による研究データマネジメントの実現のための3つのアプローチ

公的資金による研究データマネジメントに求められること（先進的データマネジメント）

- 管理対象データの範囲の特定
- メタデータの付与（課題番号・課題名称、管理者、公開・共有の区分など）
- 機関リポジトリ等への管理対象データの収載
- 研究データの管理・利活用の実施状況に関する評価体系への導入

研究開発を行う機関

・データポリシーの策定
（国立大学・大学共同利用機関法人・国立研究開発法人においては2025年まで）
・大学ファンドや地域中核・特色ある研究大学総合振興パッケージ等の大学支援策との連携（方策を2022年度中に検討）

公募型の研究資金

・ムーンショット、次期SIP等の横展開
・全ての公募型の研究費の新規公募分に導入（2023年度まで）

研究分野

・マテリアル、バイオ、防災等の研究分野からの横展開

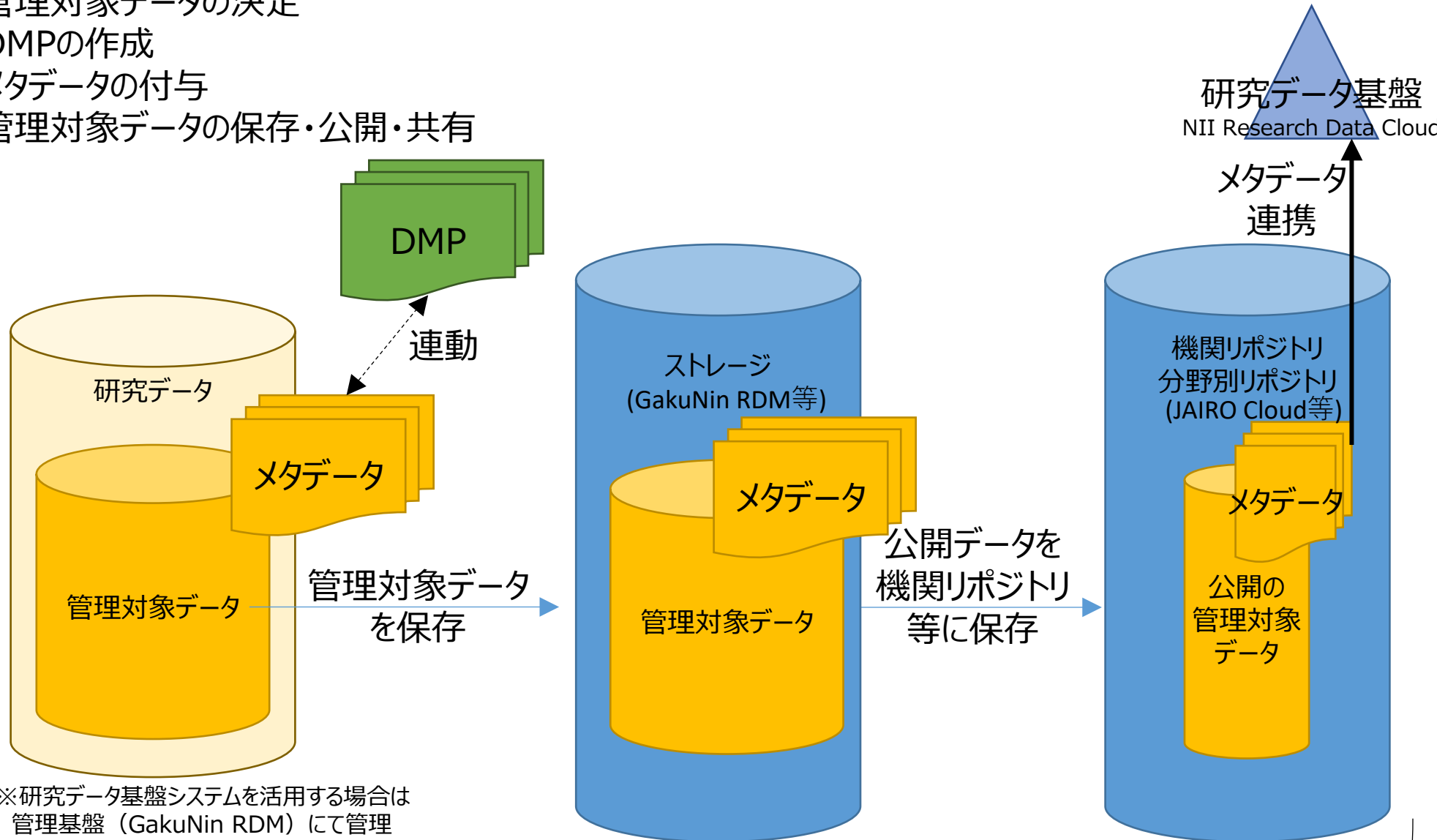
中核的なプラットフォームである研究データ基盤システム（NII-RDC）においてメタデータを検索可能な体制を構築

産学官のユーザーが研究データの所在を検索可能とし、管理者より入手できる体制を構築（第6期基本計画の計画期間2025年度までに）

* オンラインでダウンロードできることが望ましいが、別的手段でも可。

1.(4) 研究者が取り組むこと

- 管理対象データの決定
- DMPの作成
- メタデータの付与
- 管理対象データの保存・公開・共有



データポリシー、公募要領、分野におけるルール等に沿って研究データの管理・利活用を実施

(参考) 研究データに関する概念整理

研究データ

公的資金による研究開発の過程で生み出される全てのデータで、電磁的な形態により管理可能なものをいう。研究ノートやメモ、実験や観測、シミュレーション等から直接得られたデータやそれを加工したデータ、論文のエビデンスとなるデータ等が含まれる。

管理対象データ

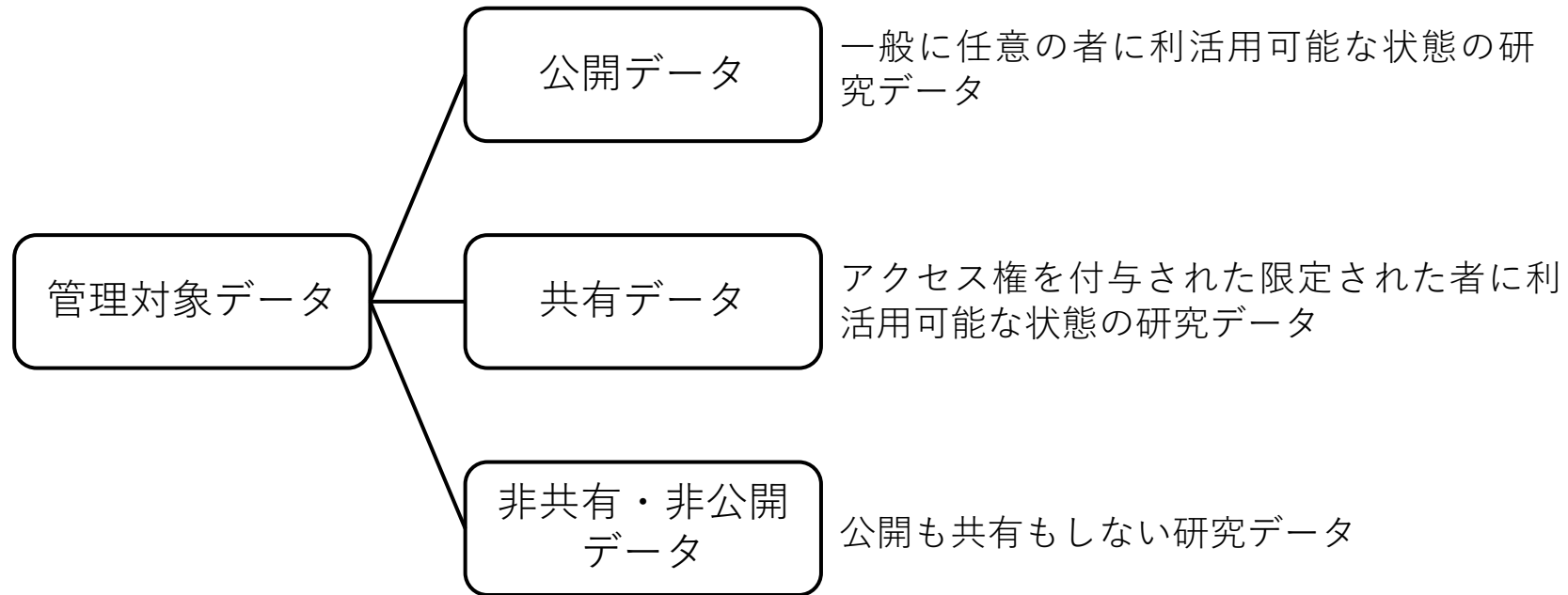
研究データのうち、研究者の所属する研究開発を行う機関や資金配分機関の基準等に基づいて、管理・利活用の対象として、研究者がその範囲を定めるものをいう。

対応

メタデータ

管理対象データを説明するための情報から構成されるデータをいう。研究データの名称、研究データの説明、研究データの管理者及びその連絡先、研究データの所在場所、研究データの保存・公開・共有の方針等の情報を含む

(参考) 管理対象データの公開及び共有の区分

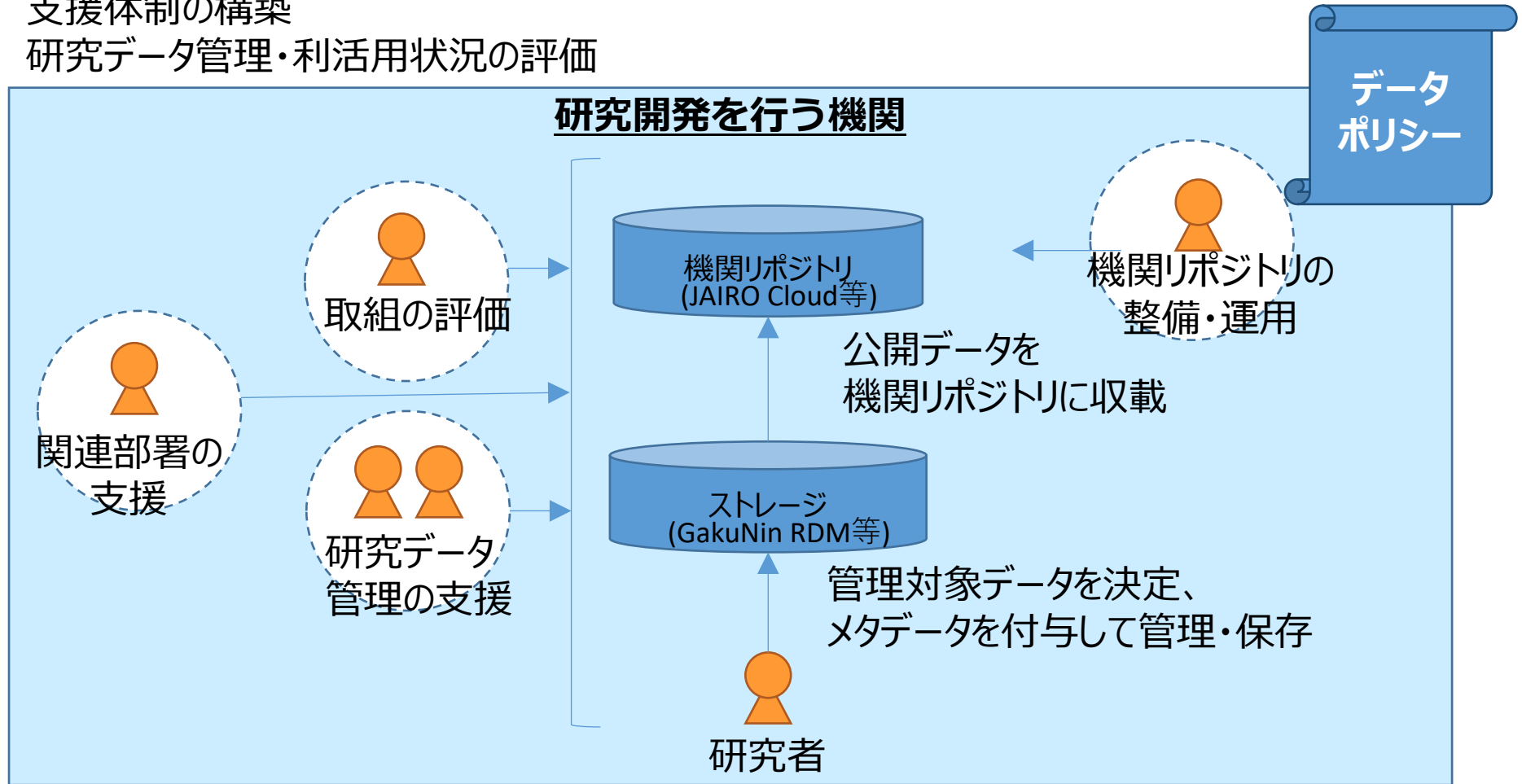


※) 「公的資金による研究データに関する基本的な考え方」から要約

- ✓ 研究分野等の特性や、大学、大学共同利用機関法人、国立研究開発法人等のデータを管理する組織の特性に配慮して、公開、共有、又は非共有・非公開の判断が行われる必要がある
- ✓ 我が国の産業競争力や科学技術・学術上の優位性を確保するために重要な情報を含む可能性があるため、個人情報、企業の秘密情報、研究の新規性、我が国の安全保障等の観点から留意すべき研究データは非公開とすることが求められる
- ✓ 産業競争力や科学技術・学術的な優位性を確保するために、公開による利活用の促進とのバランスを考慮しつつ、適切なエンバーゴ（時限非公開）期間を設定することも想定される
- ✓ 関係諸法令に従うとともに、データの取り扱いに関する各国の国内法及びEU規則並びにデータ管理の原則であるFAIR原則等の国際的な規則や慣行等との整合性に十分留意する必要がある

1.(5) 研究開発を行う機関で取り組むこと

- データポリシーの策定
- 機関リポジトリの整備・運用
- 支援体制の構築
- 研究データ管理・利活用状況の評価



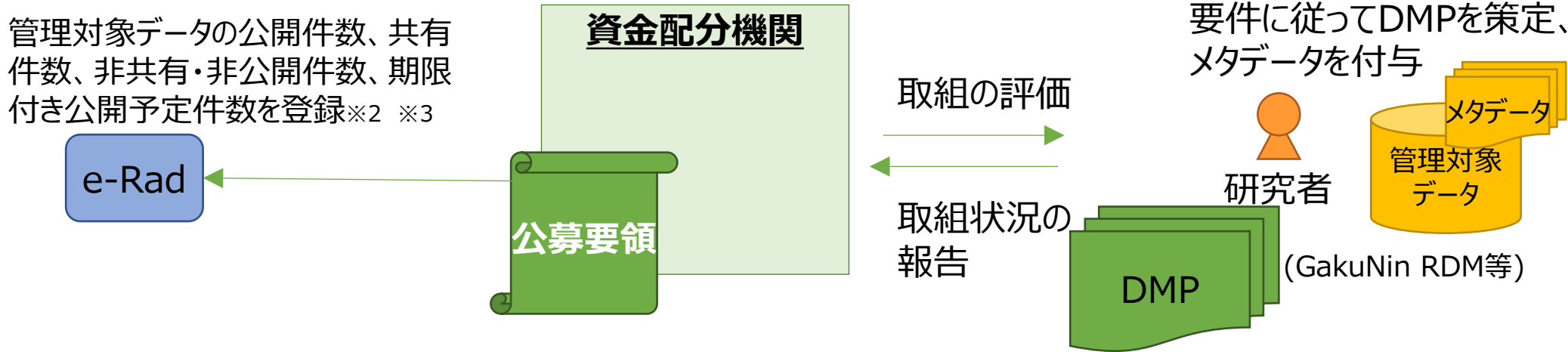
※データポリシーは機関としての研究データの管理・利活用の方針を示すもの。

※機関リポジトリを保有しない機関では、研究データ基盤NII Research Data Cloudの公開基盤JAIRO Cloudの利用、検索基盤CiNii Researchと連携した分野別リポジトリ・汎用リポジトリへのデータ収載を推奨する。これらの基盤を用いても対応が難しい場合には、各機関のシステムの活用により対応する。（これらと研究データ基盤システムとの連携は別途進める）。

1.(6) 資金配分機関で取り組むこと

- DMP項目・メタデータ項目の決定
- 取組状況の評価体系への導入※1
- 取組状況のe-Radへの登録

※公募型の研究資金の新規公募分において、DMP及びこれと連動したメタデータの付与を行う仕組みの導入率を2023年度までに100%となることを目標とする。（競争的研究費制度に含まれるものを対象とする）



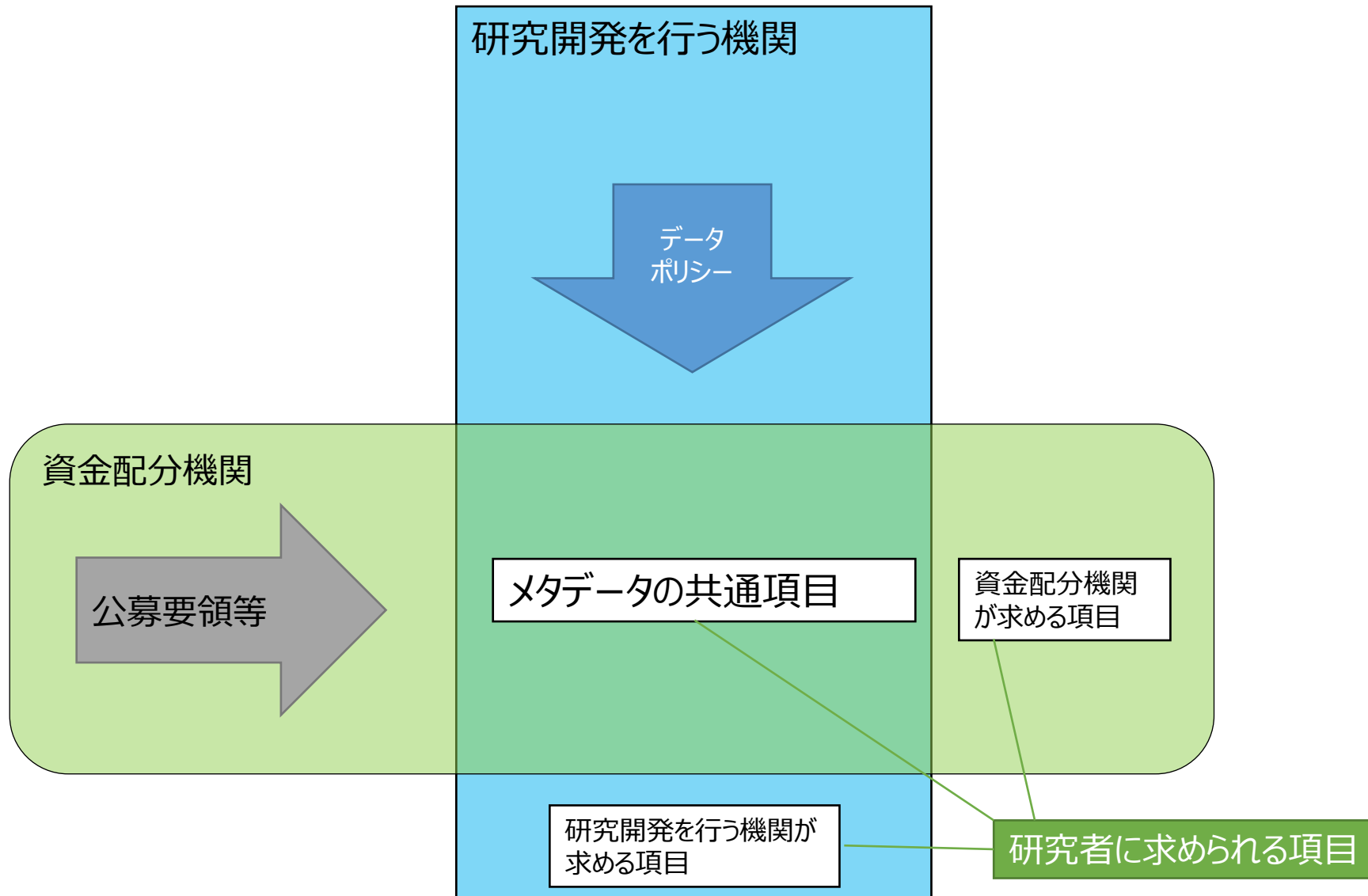
※1 単純に、公開や共有の多寡を評価するのではなく、オープン・アンド・クローズ戦略に基づいて適切に研究データの管理・利活用を行ったか、そのマネジメントについて評価する必要があることに留意する。

※2 研究開発を行う機関や研究者自身がe-Radへ登録を行う場合も含む。

※3 e-Radの改修については今後関係機関と調整。

※公開する管理対象データの収載場所については、機関リポジトリJAIRO Cloud（研究データ基盤システムの公開基盤）の利用、または、研究データ基盤システムの検索基盤CiNii Researchと連携した分野別リポジトリ・汎用リポジトリへのデータの収載を推奨する。なお、これによっても対応が難しい場合には、各機関のシステムの活用等により対応する。（これらと研究データ基盤システムとの連携は別途進める）

(参考) メタデータに関する機関と公募型の研究資金の関係



2. 実施状況（統合イノベーション戦略2024）

【第6期科学技術・イノベーション基本計画における主要指標】

- 機関リポジトリを有する全ての国立大学・大学共同利用機関法人・国立研究開発法人におけるデータポリシーの策定率（目標：2025年までに100%）
 - 国立大学：26機関 ※R5年度学術情報基盤実態調査結果より
 - 大学共同利用機関法人：4/4 法人・機関
 - 国立研究開発法人：24/24 法人・機関
- 公募型の研究資金の新規公募分における、DMP及びこれと連動したメタデータの付与を行う仕組みの導入率（目標：2023年度までに100%）
 - 2023年度末時点での競争的研究費制度における仕組みの導入率78%（141件のうち110制度（一部導入済み 13制度を含む））

【参考指標】

- 国内における機関リポジトリの構築数：858 個（2023年度）
- 研究データ公開の経験のある研究者割合：50.1%（2022年度）
- プレプリント公開の経験のある研究者割合：29.5%（2022年度）

【基本計画に基づき検討を行う指標】

- 2020年度に実施した試行的取組をベースとして、DXによる研究活動の変化等に関する新たな分析手法・指標の開発を行い、2021年度以降、その高度化とモニタリングを実施する。【文】

(参考) 統合イノベーション戦略2024

基本計画における具体的な取組	実施状況・現状分析	今後の取組方針
<p>○研究データの管理・利活用のための我が国の中核的なプラットフォームとして2020年度に本格運用を開始した研究データ基盤システム（NII Research Data Cloud）の普及・広報と必要な改良を引き続き進める。また、公的資金により得られた研究データについて、産学官における幅広い利活用を図るため、2023年度までに体系</p>	<ul style="list-style-type: none"> ・「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」を通じて、全国的な研究データ基盤の高度化及び利活用に向けた環境整備支援、大学の研究データマネジメントに係る体制・ルール整備支援を実施。 ・研究データ基盤の利用者数の増加傾向を分析し、徴収開始時期を2027年度以降に仮定した上で、受益者負担モデルについて検討開始。 ・改修に合わせ、e-Radでの実績報告時のメタデータ件数登録機能を実装。 	<ul style="list-style-type: none"> ・引き続き、全国的な研究データ基盤の利活用を推進するほか、大学の研究データマネジメントに係る体制・ルール整備を支援。【<u>科技</u>、<u>文</u>、関係府省】 ・機関、もしくは研究者が支払い可能な料金設定、かつ徴収開始後に収支均衡となる期間等について、関係者間で検討。【<u>科技</u>、<u>文</u>、関係府省】 ・メタデータ件数登録機能を運用。【<u>科技</u>】
<p>的なメタデータ¹³⁶の付与を進め、同年度以降、研究データ基盤システム上でこれらのメタデータを検索可能な体制を構築する。さらに、メタデータをEBPMに活用するため、e-Radの改修に合わせて、相互運用性を確保する。研究データ基盤システムについて、持続的な運営体制の確保に向け2022年度までに方策を検討する。【<u>科技</u>、<u>文</u>、関係府省】</p>	<ul style="list-style-type: none"> ・「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」における取組の具体化・周知のため、「研究データ2023」を策定し、関係府省への周知や事例の共有等を実施。 	<ul style="list-style-type: none"> ・研究データの管理・利活用に関する事例の収集を引き続き実施し、「研究データ2023」の更新を行う等により、関係府省での取組の具体化や周知を推進。【<u>科技</u>】
<p>○公的資金により得られた研究データの機関における管理・利活用を図るため、大学、大学共同利用機関法人、国立研究開発法人等の研究開発を行う機関は、データポリシーの策定を行うとともに、機関リポジトリへの研究データの取載を進める。あわせて、研究データ基盤システム上で検索可能とするため、研究データへのメタデータの付与を進める。【<u>科技</u>、<u>文</u>、関係府省】</p>	<ul style="list-style-type: none"> ・大学等研究機関における研究データポリシーの策定及び策定に向けた機関内での検討が各地で加速。 ・「地域中核・特色ある研究大学総合振興パッケージ」の関連事業において、大学に対する支援施策等と連動し、大学における研究データの管理・利活用に係る支援体制の整備を実施。 	<ul style="list-style-type: none"> ・引き続き、大学等研究機関における研究データポリシーの策定、及び実効力のある研究データの利活用に向けた環境整備等を進める。【<u>文</u>】 ・引き続き、大学における研究データの管理・利活用に係る支援体制の整備を推進。【<u>科技</u>、<u>文</u>、関係府省】

(参考) 統合イノベーション戦略2024

基本計画における具体的な取組	実施状況・現状分析	今後の取組方針
<p>○公募型の研究資金の全ての新規公募分について、研究データの管理・活用を図るため、データマネジメントプラン(DMP)及びこれと連動したメタデータの付与を行う仕組みを2023年度までに導入する。次期SIPにおいても同様に、DMPの策定とメタデータの付与を実施することとする。 【<u>科技</u>、<u>文</u>、<u>関係府省</u>】</p>	<ul style="list-style-type: none"> ・ムーンショット型研究開発制度における研究データの保存、共有、公開の状況を自己評価及び外部評価で確認。 ・SIP第3期において、DMPに基づくデータマネジメントを導入。 ・健康・医療データ利活用基盤協議会において、AMEDのデータ利活用プラットフォームを用いたデータ連携、同意の在り方を引き続き整理。 ・AMEDにおいて、適切なデータ利活用を促進するための各種ポリシー・ガイドラインを作成し、AMEDホームページにて公開。 ・AMEDのデータ利活用プラットフォームの整備を実施し、2024年3月から一般利用受付を開始。 <p>・DMP及びこれと連動したメタデータ付与を行う仕組みを導入した制度は78%（2023年度末時点での競争的研究費制度141件のうち110制度（一部導入済み13制度を含む））。</p>	<ul style="list-style-type: none"> ・引き続き、ムーンショット型研究開発制度における先進的データマネジメントの実施状況の検証を行う。【<u>科技</u>】 ・引き続き、DMPに基づくデータマネジメントを推進。【<u>科技</u>】 ・AMEDが支援した研究開発によって得られたデータを産学官の研究開発で活用するため、複数のデータベース等を連携し、ゲノム情報等から抽出されるメタデータを用いた横断検索機能を有するとともに、産業界も含めた研究開発にデータを扱う場(セキュリティが担保された Visiting 利用環境)を広く提供するAMEDのデータ利活用プラットフォームにおいて、ゲノムデータ以外の幅広い研究開発データの連携を順次開始。【<u>健康医療</u>、<u>文</u>、<u>厚</u>、<u>経</u>】 ・2023年度までに未導入の公募型の研究資金を所管する関係府省における、DMP及びこれと連動したメタデータの付与を行う仕組みの導入を着実に推進。【<u>科技</u>、<u>文</u>、<u>関係府省</u>】
<p>○研究データ基盤システムと内閣府が実施する研究開発課題(SIP等)で構築する分野ごとデータ連携基盤との間で、相互にデータの利活用を図るための仕組みを2023年度中に構築する。【<u>科技</u>、<u>文</u>】</p>	<ul style="list-style-type: none"> ・SIP第3期における課題間データ連携の現状分析、仮説設定を実施。 ・「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」において、全国的な研究データ基盤及び活用環境整備を実施。 	<ul style="list-style-type: none"> ・SIP第3期における課題間データ連携の仮説検証を推進。【<u>科技</u>】 ・引き続き、全国的な研究データ基盤及び当該基盤の活用環境整備を推進。【<u>文</u>】

(参考) 統合イノベーション戦略2024

基本計画における具体的な取組	実施状況・現状分析	今後の取組方針
<p>○研究者の研究データ管理・利活用を促進するため、例えば、データ・キュレーター、図書館職員、URA、研究の第一線から退いたシニア人材、企業等において研究関連業務に携わってきた人材、自らの研究活動に資する場合にはポストドク等の参画や、図書館のデジタル転換等の取組について、2022年度までにその方向性を定める。【<u>科技</u>、<u>文</u>、関係府省】</p>	<ul style="list-style-type: none"> ・「『2030 デジタル・ライブラリー』推進に関する検討会」を設置し、「デジタル・ライブラリー」の実現に向け、大学図書館が検討すべき取組の方向性を具体化するロードマップを作成。「デジタル・ライブラリー」の実現における課題の洗い出しの調査を実施。 ・「地域中核・特色ある研究大学総合振興パッケージ」関連事業において大学に対する支援施策等と連動し、研究データの管理・利活用に係る支援体制整備を実施。 	<ul style="list-style-type: none"> ・「『2030 デジタル・ライブラリー』推進に関する検討会」において、「デジタル・ライブラリー」の実現における課題の洗い出しに向けた調査の結果を踏まえ、優先的に取り組むべき課題を整理。また、ロードマップの修正も適宜検討。【<u>科技</u>、<u>文</u>、関係府省】 ・引き続き、大学における研究データの管理・利活用に係る支援体制の整備を推進。(再掲)【<u>科技</u>、<u>文</u>、関係府省】
<p>○自由で開かれた研究活動を尊重し、我が国と価値観を共有する国・地域・国際機関等（EU、G7、OECD等）との間で、研究データの管理・利活用に関する連携を進める。我が国の研究データ基盤システムとこれに相当する取組との国際連携を図り、研究データの管理・利活用に関する国際的な相互運用性を高めることにより、本計画期間中に、グローバルプラットフォームの構築を目指す。【<u>科技</u>、<u>文</u>】</p>	<ul style="list-style-type: none"> ・G7 仙台科学技術大臣会合の共同声明及びG7 広島首脳コミュニケにおいて、オープンサイエンスの推進が明記され、これらを基にオープンサイエンスWGの活動を実施。 ・EUのEOSC（European Open Science Cloud）と我が国の研究データ基盤システム間での相互運用性を目的とする共同研究および連携プロジェクトを立ち上げ、検討を開始。 	<ul style="list-style-type: none"> ・G7 仙台科学技術大臣会合を踏まえたオープンサイエンスWGの活動を具体化。【<u>科技</u>、<u>文</u>】 ・EOSCと研究データ基盤システム間の相互運用性や欧州におけるオープンサイエンスモニター（Open Science Monitor）等との連携の実現可能性を検討。【<u>科技</u>、<u>文</u>】
<p>○研究データの管理・利活用に関する取組を更に促す観点から、2022年までに、これらの取組の状況を、研究者、プログラム、機関等の評価体系に導入する。【<u>科技</u>、関係府省】</p>	<ul style="list-style-type: none"> ・一部関係府省の事業において、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を実施。 	<ul style="list-style-type: none"> ・引き続き、関係府省にて研究者、プログラム、機関等の評価体系への導入を検討。【<u>科技</u>、関係府省】

3. メタデータ説明書について

- 先行的な先進的データマネジメント導入事例であるムーンショット型研究開発制度のメタデータ項目を説明するために「ムーンショット型研究開発制度におけるメタデータ説明書」を作成。
- 「ムーンショット型研究開発制度におけるメタデータ説明書」の運用に基づき、「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 関係府省申合せ）に関する補足説明のために、「公的資金による研究データの管理・利活用に関するメタデータ説明書」を作成。

ムーンショット型研究開発制度におけるメタデータ説明書（第3.1版）



https://www8.cao.go.jp/cstp/ms_metadatainstructions.pdf

公的資金による研究データの管理・利活用に関するメタデータ説明書（第1.0版）



<https://www8.cao.go.jp/cstp/metadatainstructions.pdf>

4. ムーンショット型研究開発制度における取組事例

4. 事例紹介 (JST・目標1)

石黒プロジェクトで収集されたデータの有効活用事例

石黒プロジェクトでは、サイバネティック・アバター (CA) システムを通して得られる多様なデータの再利用に取り組んでおり、計24件の体系的なデータ収集と整備を行っている。14件がプロジェクト内で共有されており、このうち2件がGitHubを通して広く公開されている。

日本語日常対話コーパス (日本語の日常的な対話データ)

日常生活、学校、旅行、健康、娯楽の5つのトピックに関する日常会話を収録した高品質なマルチターン対話データセット。すべての対話は基本的な語彙と語順で標準的な日本語で書かれている。

公開場所：GitHub (<https://github.com/jqk09a/japanese-daily-dialogue>)

公開時期：2023年5月17日

アクセス数：195件/2週間 (2024年8月15日～8月28日に集計)

活用事例：2023年度言語処理学会の予稿の google scholar の被引用数は6件、それ以外に検索でわかった事例が4件 (2024年8月時点)

BPersonaChat (日本語と英語の対訳がセットとなった対話データ)

英語多言語チャットコーパスPersona-chatと日本語多言語チャットコーパスJPersona-chatに基づく評価データセット。各チャットは、人工的なペルソナを想定した2人のクラウドワーカーの間で行われ、発言者は、自己紹介、趣味、その他を含むがこれらに限定されていない。

公開場所：GitHub (<https://github.com/cl-tohoku/BPersona-chat>)

公開時期：2023年1月12日

利用件数：14件 (2024年8月時点。日常的なアクセス件数は不明)

活用事例：対話システムに翻訳アシスタントを導入するための調査で利用

4. 事例紹介 (JST・目標1)

金井PM「身体的能力と知覚能力の拡張による身体の制約からの解放」

非侵襲BMI*操作中の頭皮脳波データ

*ブレイン・マシン・インターフェース (BMI) : 脳の運動指令を脳波から読み取り、コンピュータによって分析することで、車椅子や家電、義手、アバターなどを直接動かす技術。

公開場所 : OpenNeuro で一般に公開

(<https://openneuro.org/datasets/ds004448/versions/1.0.0>)

最終更新日 : 2023-01-26

アクセス数 : 論文へのアクセス数は3114件 (2024年8月5日時点)

※本データセット自体へのアクセス件数はカウント不可であった。

※参考までに、本データセットに関する論文のアクセス件数は上記のとおり。

方法 : 頭部全体をカバーする 128 チャンネルを含む高密度頭皮 EEG 装置で計測

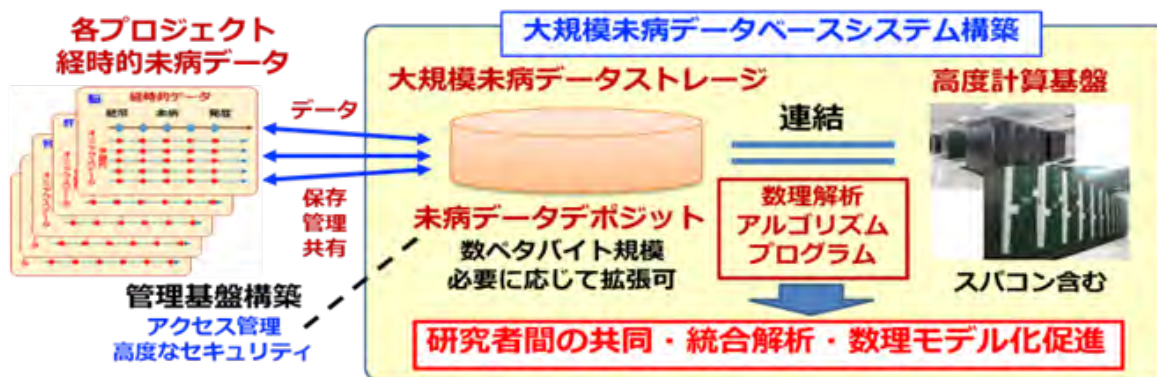
活用事例 : 安静時頭皮脳波 (EEG) データから個々のアルファ周波数 (IAF) を決定するアルゴリズムの有効性を上記の大規模脳波データを用いて確認した。これにより、脳波における個々のIAFを迅速かつ確実に識別することが可能となり、アバターを介した神経通信および脳コンピュータインターフェースの使用に対しての貢献が期待される。

Seitaro Iwama, Junichi Ushiba. IEEE Transactions on Neural Systems and Rehabilitation Engineering (Volume: 32, 915 - 922). 12 February 2024

4. 事例紹介 (JST・目標2)

2050年までに、超早期に疾患の予測・予防をすることができる社会を実現

- 【目的】各プロジェクトで創出されたデータを共有し、円滑・高速な統合解析、数理モデル化を行うため、大規模データベースシステムを構築中。

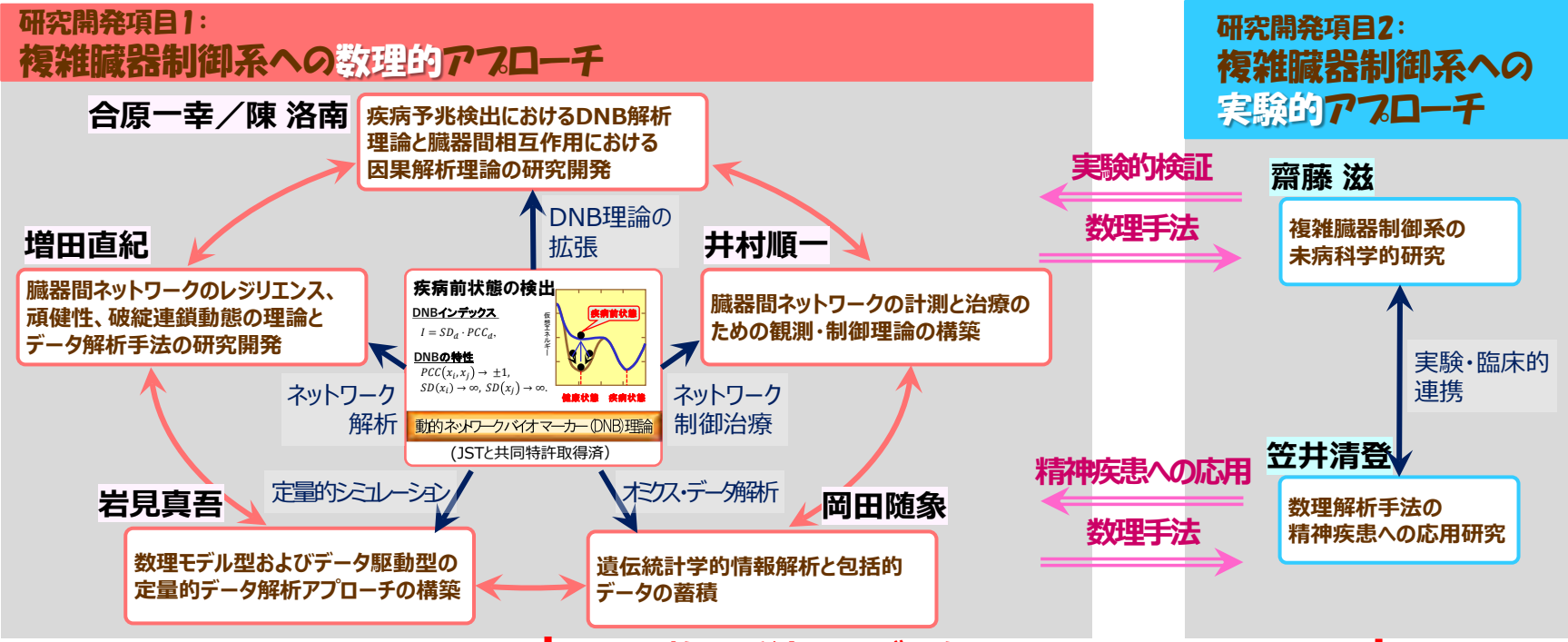


NII (国立情報学研究所) GakuNin RDM

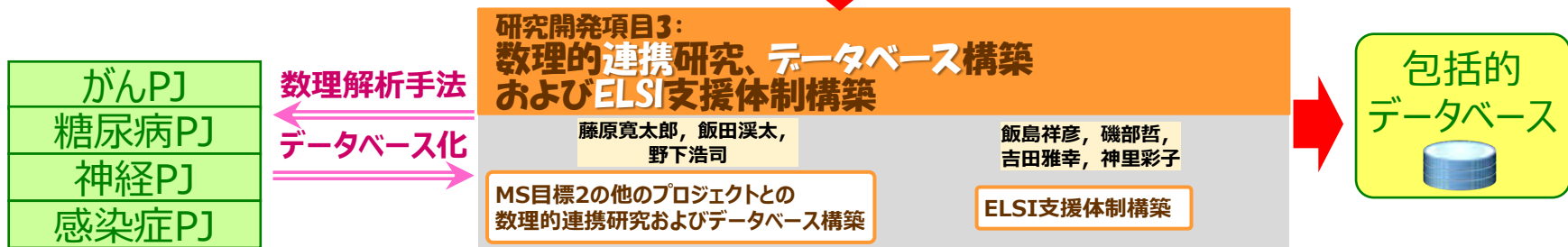
- 【DBシステム】NII (国立情報学研究所) のGakuNin RDMを利用。
- 【体制】全プロジェクト横断で、数理データ連絡会議、データベース作業部会を構築し、最適なDB運用に向けて議論し、データフォーマット、メタデータ設計、規程整備、データベース利活用規程等を策定し推進。DBマネジメントやELSIの専門的な側面支援を目的に、データベースマネジメント支援チーム、ELSI対応チームを構築。
- 【管理データ】各プロジェクト経時的未病データ (Bulk RNA-seq、一細胞RNA-seq、空間的遺伝子発現、ゲノム・エピゲノム等)
- 【共有範囲／将来構想】：当初は目標2内の共有を実施。将来的には、本格的な未病社会の構築に向けた統合的な超早期の疾患の予測・予防の研究において、国際的な未病データ基盤の礎となることを想定。

4. 事例紹介 (JST・目標2)

合原一幸PM 「複雑臓器制御系の数理的包括理解と超早期精密早期精密医療への挑戦」



数理手法 と データ



4. 事例紹介 (JST・目標3)

菅野プロジェクトのAI学習データの有効活用事例

【研究概要】

菅野プロジェクト「一人に一台一生寄り添うスマートロボット」におけるAI領域では、複数の基本タスクの動作学習に関する研究を実施している。このタスクの学習では、早稲田大学発の深層予測学習（Deep Predictive Learning）を用いており、学習データの蓄積を行っている。

【データ公開の狙い】

オープン戦略に基づいて、深層予測学習モデルのデータを公開することで、他プロジェクト（異なるロボット）への利活用と連携を図る。

【深層予測学習モデルのオープンソースライブラリ】

深層予測学習モデルのソースプログラムと3種類のデータをセットで公開している。

■ 3種類のデータ

- 菅野プロジェクトで開発中の多自由度スマートロボット（AIREC）にて物体を把持したデータ1種
- 市販の安価なロボットアーム（OpenManipulator）の物体把持データ1種
- シミュレータのデータ1種

■ 公開場所：GitHub (<https://ogata-lab.github.io/eipl-docs/top/>)

■ 公開時期：2023年6月28日

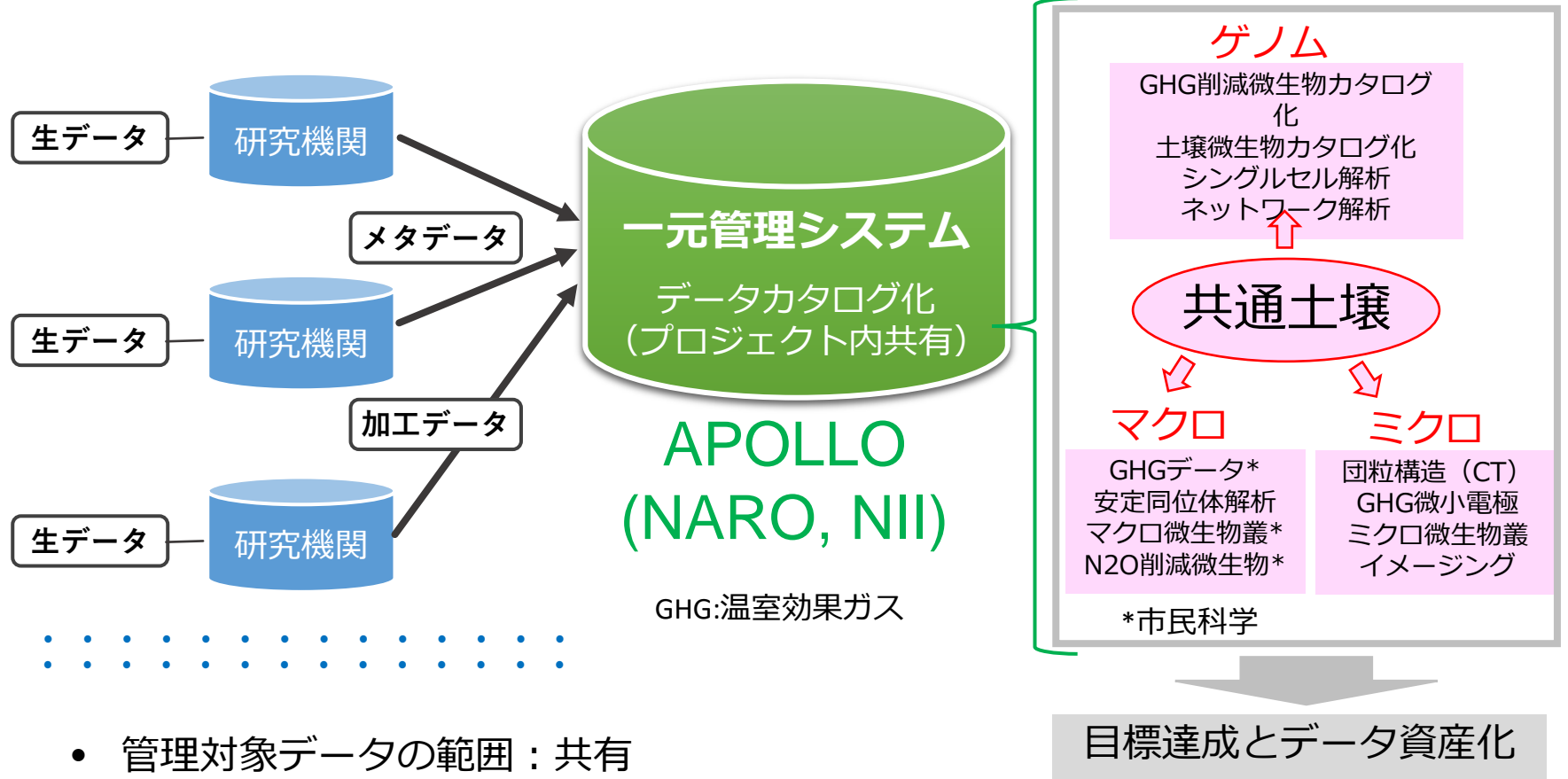
■ アクセス数：1731(2024年8月5日時点)。日本1291、アメリカ184、中国124、韓国、英国など。

4. 事例紹介 (NEDO・目標4)

南澤 究PM 「資源循環の最適化による農地由来の温室効果ガスの排出削減」

取組事項

- プロジェクト内のGHG削減土壌メタデータを一元管理するシステムを構築
- 蓄積されたデータをプロジェクト内で活用する取り組みを実施

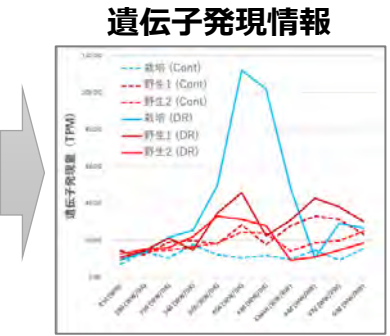
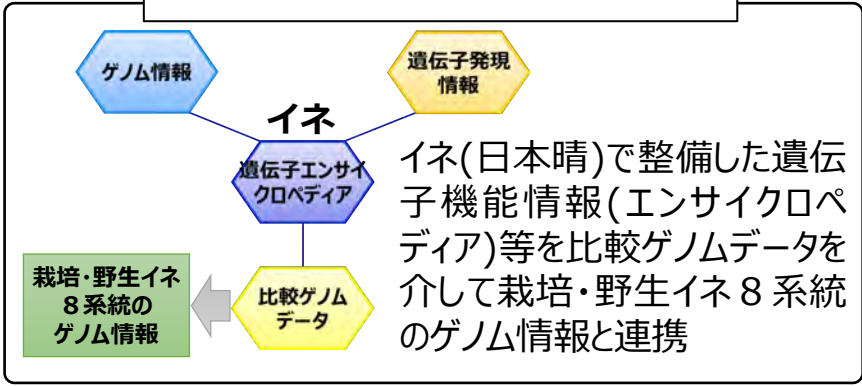


- 管理対象データの範囲：共有
- 公開、共有、非公開・非共有の区分の基準：共有

4. 事例紹介 (BRAIN・目標5)

藤原PM「イネ版作物横断型データ基盤の開発」

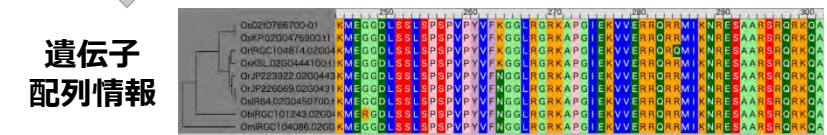
イネ版作物横断型データ基盤の概要



遺伝子機能情報

Reference: NP, KP, Kasahiri, IRG104814, JP223922, JP226069, IR84, IRG101543, IRG104086
Species: NP, KP, Kasahiri, IRG104814, JP223922, JP226069, IR84, IRG101543, IRG104086
Method: Orthologous (OrthoFinder), Synteny (MCSort), Bush

Ortholog Table	Reference	Information	Description	Symbol	NP	KP	Kasahiri	IRG104814	JP223922	JP226069	IR84	IRG101543	IRG104086
Os02g0768700	Os02g0768700.1	NP223922.020043	5ZP3 Os2P3										



当該データ基盤開発の詳細

- 遺伝子エンサイクロペディア**: イネ有用遺伝子637個をカタログ化しRAP-DBから公開
- ゲノム情報**: 栽培、野生イネ8系統のリファレンスゲノム配列、遺伝子アノテーション情報を整備
- 遺伝子発現情報**: 栽培、野生イネ8系統594サンプルの遺伝子発現プロファイルを整備
- 比較ゲノムデータ**: 日本晴をベースに、栽培・野生イネ8系統の遺伝子を相互に対応付け

各データを可視化するアプリケーションを開発しウェブシステムとして統合

2024年度よりイネアノテーションデータベース(RAP-DB: <https://rapdb.dna.affrc.go.jp/>)上で随時公開

<活用事例>
乾燥耐性を持つ野生イネ系統に特徴的な遺伝子発現パターンを見出し、イネの品集改良を迅速化

4. 事例紹介 (BRAIN・目標5)

竹山PM 「土壌微生物叢アトラスに基づいた環境制御による循環型協生農業プラットフォーム構築」

【取組内容】

- 研究データは、研究機関ごとの管理からプロジェクト全体での共有管理への移行を推進中。
- 「**土壌微生物アトラス**」と「**農業デジタルツイン**」を基軸とするデータマネジメントによって、**地域・日本・地球の健康**に貢献する研究成果を追求。

<土壌微生物アトラス> 土壌微生物の特徴(系統・遺伝子)を土壌の性質ごとに体系化し、土壌微生物を中心にした地図帳 (アトラス) を構築



データ収集

- 細菌叢データ
- 細菌ゲノムデータ
- ラマン分光データ
- 土壌・環境・作物データ etc.



データ分析

- 有用微生物の共起ネットワーク解析
- RNA-seq 解析
- 代謝物解析 etc.



データ活用

- 土壌健康度の指標
- 微生物資材
- 食味と生産性を兼ね備えたダイズ育種 etc.

プロジェクト完了時の全データは、～ 1 PB (千兆バイト) 規模のデータ量を想定

<農業デジタルツイン> 土壌微生物を含めた農業のデジタル管理を中心にワンヘルスを実現



データ収集

- 収集システム
- データ蓄積・加工



データ分析

- モデル構築
- 予測結果導出



データ活用

- 分析ツール提供
- API提供

土壌分析関連機能

収量予測



品質予測



環境負荷予測



サプライチェーン関連機能

品種別需給管理



直接取引市場 (国内/海外)



スマートフード促進
コミュニティ管理



大規模圃場試験 (北海道～九州の6地点) と全国圃場 (33道府県59圃場) で取得した、土壌微生物アトラスデータを含む数万を超える測定項目のマルチモーダルデータ (マルチオミクス) で学習された統合モデルを搭載

- **管理対象データ**：論文のバックデータを必須としつつ、可能な範囲でデータの共有をプロジェクト内に求めている。
- **公開、共有、非共有・非公開の区分**：業界を発展させるために必要なデータは、可能な限り公開・共有。知財の保護等に係るものは、非共有・非公開。

出典：BRAIN提供

4. 事例紹介 (BRAIN・目標5)

由良 敬PM 「地球規模の食料問題の解決と人類の宇宙進出に向けた昆虫が支える循環型食料生産システムの開発」

【取組内容】

- 研究データは研究者が所属する各研究機関で管理しているが、プロジェクト内で研究データを参照（共有）できるシステムを開発中。
- バイオインフォマティクスの観点から、多様な研究データを活用・分析することで研究成果を生み出すことを一つの目標としている。



出典：「ムーンショット型農林水産研究開発事業『地球規模の食料問題の解決と人類の宇宙進出に向けた昆虫が支える循環型食料生産システムの開発』」 (<https://if3-moonshot.org/rd/subproject/>)

- 管理対象データ：論文のバックデータを必須としつつ、可能な範囲でデータの共有をプロジェクト内に求めている。
- 公開、共有、非共有・非公開の区分：業界を発展させるために必要なデータは可能な限り公開・共有。また、知財の保護等に係るものは非共有・非公開。

4. 事例紹介 (AMED・目標7)

村上 正晃PM 「病気につながる血管周囲の微小炎症を標的とする量子技術、ニューロモデュレーション医療による未病時治療法の開発」


取組事項

- プロジェクト内の共同研究データを一元管理するシステムを構築中。
(村上PMらが管理するデータサーバーを利活用)
- 研究データの共有により、データ解析を行うことで、量子技術による超高感度解析、ニューロモジュレーション医療を実現することを目標に研究を実施。


1. 病気の芽を診る技術関連データ

最先端の量子技術


次世代デバイスによる
検出・解析と情報発信





・ IL-6アンプ因子



・ 自己抗体
自己反応性細胞




・ PET・MRI



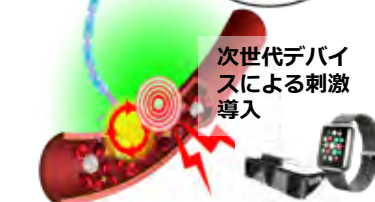
Bottom row of icons: 1. Microscopic view of cells, 2. Y-shaped antibody, 3. Cell with receptor, 4. Hand holding a device, 5. Heart with Ca²⁺ and K⁺ ions.

2. 病気の芽を摘む技術関連データ

ニューロモジュレーション戦略 + 自己反応性細胞
IL-6アンプ
阻害



次世代デバイスによる刺激
導入



- 管理対象データの範囲
個人情報
個人情報は匿名化を行うことを徹底
量子計測デバイス関連データ、大容量画像データ、遺伝子発現関連データ
生理・行動情報に関するデータなどが対象
- 公開、共有、非公開・非共有の区分の基準
 - 実験データの取得から解析までが一つのグループで完結しない場合はデータの公開・共有を進める

5. 関係府省における取組事例

5. 事例紹介（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

背景・課題

- ポストコロナの原動力として「デジタル」「AI」が最重要視され、データ駆動型研究やAI等の活用による大量の研究データ分析が世界的に進展している中、大規模かつ高品質なデータの利活用の推進を、様々な分野・機関を超えて進めていくことが鍵。
- 我が国でもオープン・アンド・クローズ戦略に基づき全国の研究者が、分野を問わず必要な研究データを互いに利活用することで、優れた研究成果とイノベーションを創出していく環境の整備が急務。
- 令和6年開催のG7科技大臣会合でも、前年に引き続きオープンサイエンスを進める旨の共同声明が出されており、研究データ利活用は世界的な潮流。

本事業で解決する課題

- ✓ 様々な研究データの利活用が、研究者の負担なく円滑に促進されるよう、研究データ基盤の高度化（他機関連携も含む）を進める。
- ✓ 適切な研究データの管理・公開、分野・機関横断的な検索といった研究データ管理・利活用が持続的に行われる仕組みを構築。
- ✓ 各研究機関が、オープンサイエンス・オープンアクセスの世界的な潮流に対応していくための体制整備にも貢献。

【G7ポロニーヤ科学技術大臣会合 共同声明】（令和6年7月9日-11日開催）

- 公的資金による学術出版物及び科学データへのオープンで公共的なアクセスを含む、科学的知識及び適切な研究成果の公平かつ責任ある普及を通じてオープン・サイエンスを拡大するため、G7メンバー及び国際的な科学コミュニティ全体の協力を促進する。

【学術論文等の即時オープンアクセスの実現に向けた基本方針】

（令和6年2月16日統合イノベーション戦略推進会議決定）

- (4) 研究成果発信のためのプラットフォームの整備・充実
- 研究成果を誰もが自由に利活用可能とするための発信手段として、研究データ基盤システム（NII Research Data Cloud）、その他のプレプリント、学術論文等の研究成果を管理・利活用するためのプラットフォームの整備・充実に対する支援を行う。

必要な取組（事業期間：令和4年度～令和8年度）

● 全国的な研究データ基盤（NII RDC※）を高度化

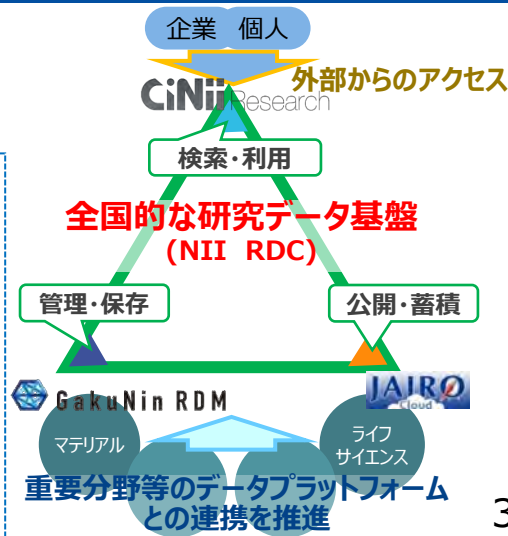
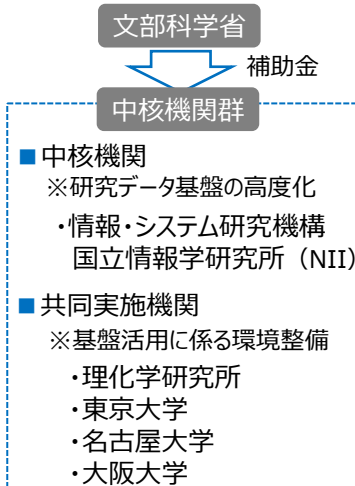
- ・研究者が研究により時間を割くことが可能となり、研究データ利活用が促進されるよう、管理データの取捨選択・メタデータ付与、データの出所・修正履歴の管理など、研究データ管理に係る関係者の作業負担を軽減するための機能等を開発

※NII-RDC（Research Data Cloud）：研究データサイクルを支える3つのシステムにより構成
管理基盤（GakuNin RDM）、公開基盤（JAIRO Cloud）、検索基盤（CiNii Research）

● 研究データ基盤の活用を促進するための環境整備

- ・全国の研究者が統一した基準でデータ管理できるよう、機械可読データの統一した記述ルールやデータ管理・公開ガイドライン整備、データマネジメント人材育成支援、各機関の研究データ基盤との連携等を実施

【事業スキーム】



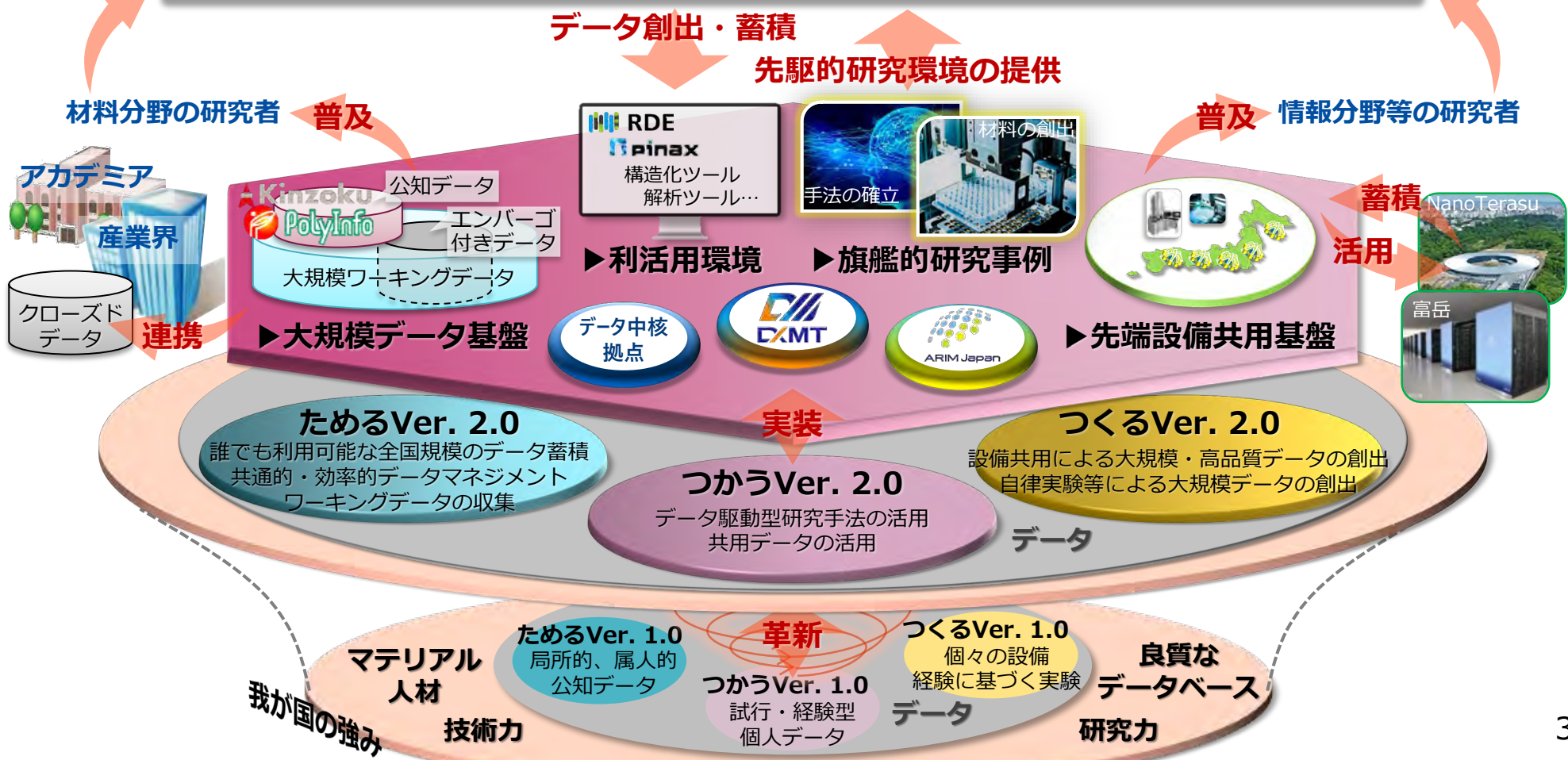
5. 事例紹介 (文部科学省) マテリアルDXプラットフォームの目指すべき姿

日本における革新的材料の創出と社会課題の解決を加速

全国の研究者によるデータ駆動型研究手法を活用した成果の創出

材料分野のプラットフォームユーザ
プラットフォームの三事業、大学・研究機関、産業界の研究者 等

情報分野等のプラットフォームユーザ
データサイエンティスト 等



5. 事例紹介 (文部科学省)

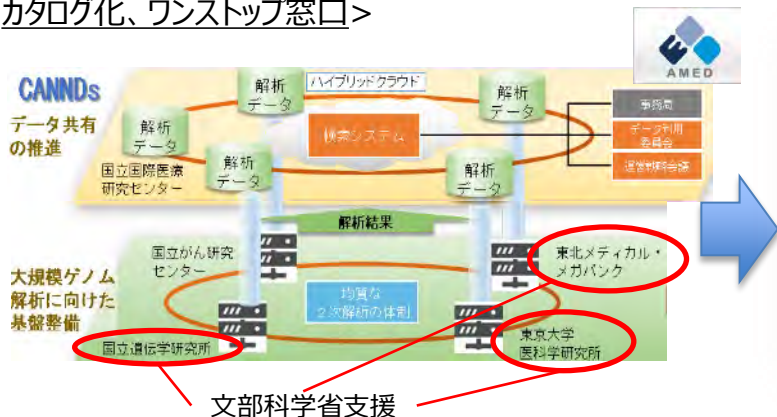
ライフサイエンス分野におけるユースケースの形成・普及

【概要】

有用な研究データの幅広い利活用の推進を目指し、AMEDを中心として府省横断的に進められているヒトゲノム情報等のデータ利活用に関する取組を推進するとともに、AIとビックデータ、実験装置等の自動化等を組み合わせた革新的な研究アプローチのユースケース創出を推進する。

健康・医療情報 ※個人情報保護に留意

- ヒトゲノム情報や臨床情報等のデータ利活用に関する取組を推進 (内閣府とりまとめ) <Visiting型計算環境、検索機能、カタログ化、ワンストップ窓口>

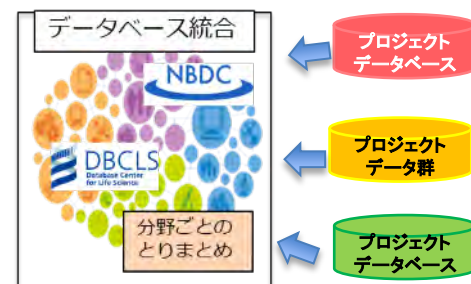


- 健康情報等の集約・匿名・仮名化 (次世代医療基盤法 (内閣府とりまとめ) 活用等)

多様な研究データを収集するとともに、ユーザーが使いやすい高品質な研究データを提供

その他ライフサイエンス関係データ

- 研究対象ごとに規格が異なるライフサイエンス分野の研究データを統合したデータベースを構築し、データ解析・利活用を推進
- AI等を活用したデータ利活用による新たな知見の創出を推進



- バイオリソースの利活用促進のため、バイオリソースの提供支援にあたりゲノム情報等の付加を推進
- 理研において、メタデータを付与したデータ (脳画像等) や解析ツールを公開

研究者、製薬企業等活用

ユーザーニーズを踏まえた更なるデータの創出

研究データ管理・統合
データ利活用の推進

研究データ
創出



クライオ電子顕微鏡

- ロボットや遠隔技術による実験の自動化
- 計測機器・技術の高度化による高解像度・時系列のデータの大量取得
- 科学研究向けの基盤モデル (生成AIモデル) を開発 (TRIP-AGIS) 【取組例】
- ✓ クライオ電子顕微鏡における試料調製の自動化、操作の遠隔化
- ✓ 理研が有する研究基盤・装置群と連携し、創薬、ワクチン関連研究や、バイオものづくりに資する分子デザイン・化学合成を推進



理研の強みを特に生かせる領域

- 多階層 (分子~個体) ・マルチモダル (オミクス、イメージング等) データの高効率計測・統合解析によるブレイクスルー
- <毛包の形成過程>
1細胞遺伝子発現解析と時系列イメージングの統合による高解像度4Dイメージング
- <ロボット実験>
ロボット技術と画像AIの統合による高効率・高品質実験



5. 事例紹介（文部科学省）

地球環境データ統合・解析プラットフォーム事業の概要

背景・課題

- 平成28年11月の「パリ協定」発効や平成30年12月の「気候変動適応法」施行等を踏まえ、科学的知見に基づく、具体的な温室効果ガスの削減取組や、気候変動の影響への適応等の対策の推進が強く求められている。
- 各国の気候変動政策等の基礎となる科学的知見を提供する気候変動に関する政府間パネル（IPCC）の活動に貢献する必要。
- 国内において、昨今の自然災害の激甚化・頻発化への対応をはじめとする気候変動対策やカーボンニュートラルに向けた取組を加速する必要。

【政策文書における記載（抄）】

<第6期科学技術・イノベーション基本計画（令和3年3月）>

・高精度な気候変動予測情報の創出や、気候変動課題の解決に貢献するため温室効果ガス等の観測データや予測情報などの地球環境ビッグデータの蓄積・利活用を推進する。

<統合イノベーション戦略2024（令和6年6月）>

・D I A Sを長期的・安定的に運用するとともに、気候変動対策の基盤となる地球環境ビッグデータの蓄積・統合・提供や、D I A Sの解析環境を活用した共同研究を促進し、データ駆動による気候変動対策に向けた研究開発を実施。

事業概要

【事業の目的・目標】

- 地球環境分野のデータプラットフォームであるデータ統合・解析システム（DIAS）の長期・安定的運用を通じて、地球環境ビッグデータ（地球観測データ、気候予測データ等）を活用した気候変動、防災等の地球規模課題の解決に貢献する研究開発や地球環境分野のデータ利活用を更に加速する。

【事業概要・イメージ】

- ・DIASの長期的・安定的な運用を通じて、大容量ストレージに地球環境ビッグデータを蓄積するとともに、統合・解析及び提供を実施
- ・地球環境ビッグデータを利活用した気候変動、防災等の地球規模課題の解決に貢献する研究開発を加速
- ・GEO（地球観測に関する政府間会合）やIPCC等を通じた国際貢献、学術研究を一層推進
- ・アジア・太平洋地域におけるデータに基づく水災害対策等を担う人材の育成に貢献
- ・より幅広い主体による共同研究等を実施するための解析環境の整備を推進



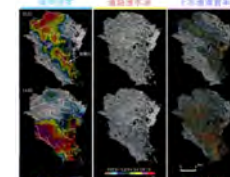
【事業スキーム】

- ✓ 支援対象機関：国立研究開発法人海洋研究開発機構
- ✓ 事業期間：令和3年度～令和12年度

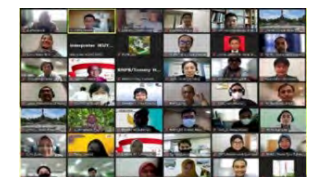


【これまでの成果】

- ✓ 国内外の幅広い分野の利用者による地球環境データの利活用が増加し、DIASの解析環境を活用した企業との有償による共同研究課題を令和5年12月から実施
- ✓ 道路や街区等の浸水状況を予測するリアルタイム浸水予測システムや台風等による洪水予測システムを開発
- ✓ 東南アジアや西アフリカを中心として、観測・予測データを活用した水災害対策に関するe-ラーニングプログラム等を用いた人材育成を実施



リアルタイム浸水予測システム



e-ラーニング・ワークショップの様子

地球環境データベース

Global Environmental Database

地球環境データの管理・公開・活用のための研究データ基盤

国立環境研究所 地球システム領域 白井知子 室長

研究内容

- 地球環境データの発信を強化するための基盤データベースとして構築、2014年から公開。
- 地上ステーション、船舶、航空機、人工衛星など、多様なプラットフォームを用いた、日本最大の温室効果ガスの観測網である地球環境モニタリング事業のデータをはじめとする多様な地球環境データの公開・活用促進に向けたサービスを提供。

特色

- DOI (デジタルオブジェクト識別子) を付与してデータ公開可能
- 提供データを機械判読可能なフォーマットに統一 (AMES)
- クイックプロット機能 (自動グラフ表示)
- リアルタイムデータ (速報値) の充実
- 複数のデータ検索手法
- 解析支援ツール (流跡線解析など)



応用例

研究データ管理システムを利用した、チームで協働しながらデータのライフサイクルを主体的に回せる仕組み作り、GEDからのスムーズなデータ公開支援を実現するデータ基盤整備を進めています。

GERDaMS 研究データ管理システム

- ◆ 研究データ管理
 - ・メタデータ作成
 - ・ライセンス付与
 - ・DOI付与
 - ・バージョン管理



GED 地球環境データベース

- ◆ データベース
 - ・データ公開
 - ・データ検索
 - ・解析支援



セールスポイント

- 地球環境データの収集・管理・公開・活用促進のためのデータ基盤の開発・運用
- データ公開やデータへのDOI付与の支援を通じたオープンサイエンス促進
- データ活用促進のため、他のデータ公開基盤との連携も積極的に進める予定