

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2023～



令和6年3月

内閣府 科学技術・イノベーション推進事務局

目次

- 研究データの管理・利活用の推進
- 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
- 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
 - ・ 研究者が取り組むこと
 - ・ 研究開発を行う機関で取り組むこと
 - ・ 資金配分機関で取り組むこと
- 実施状況（統合イノベーション戦略2023）
- ムーンショット型研究開発制度や関係府省での取り組み事例

公的資金による研究データの管理・利活用の推進

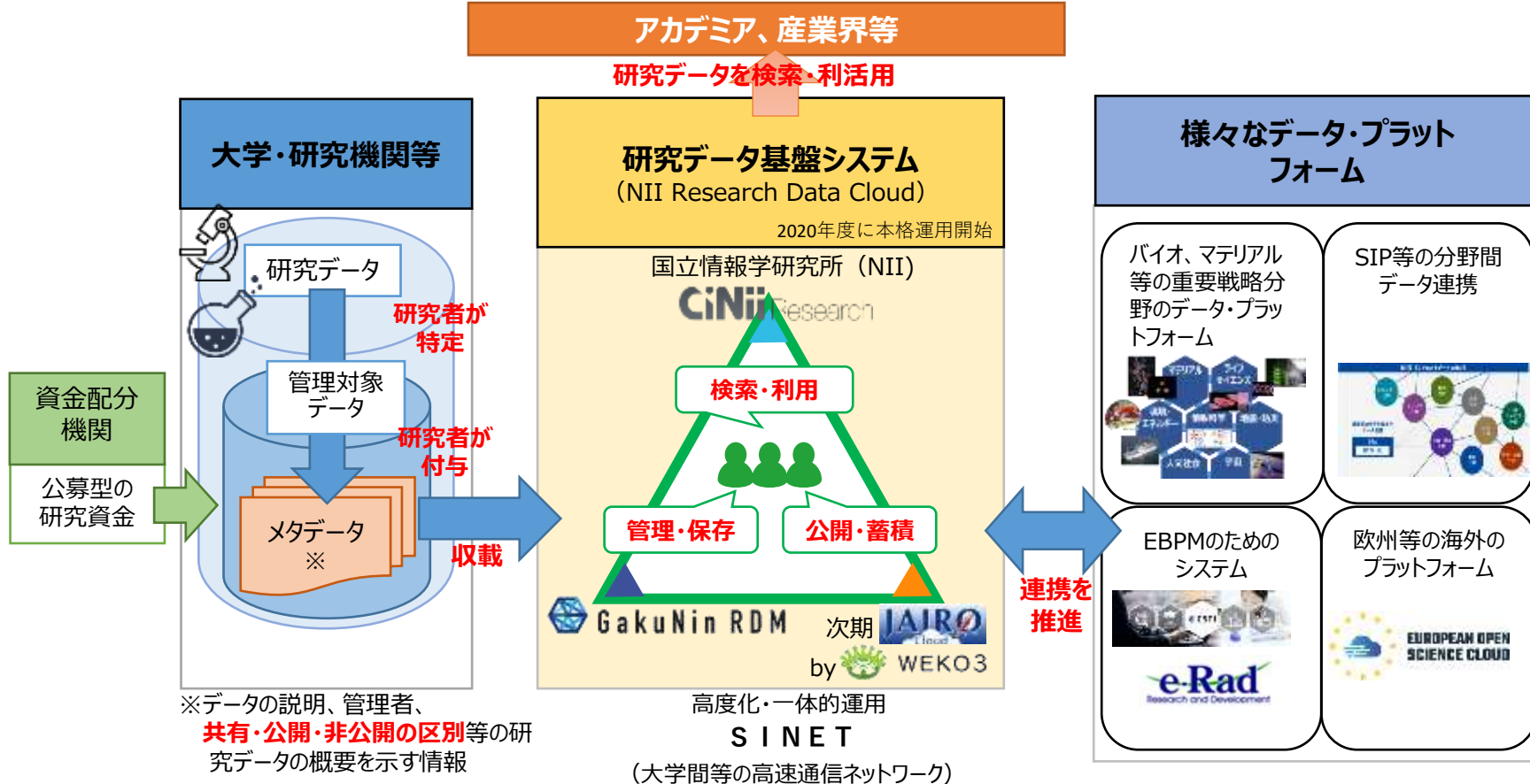
【背景】

- 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2023（2023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**



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

<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）

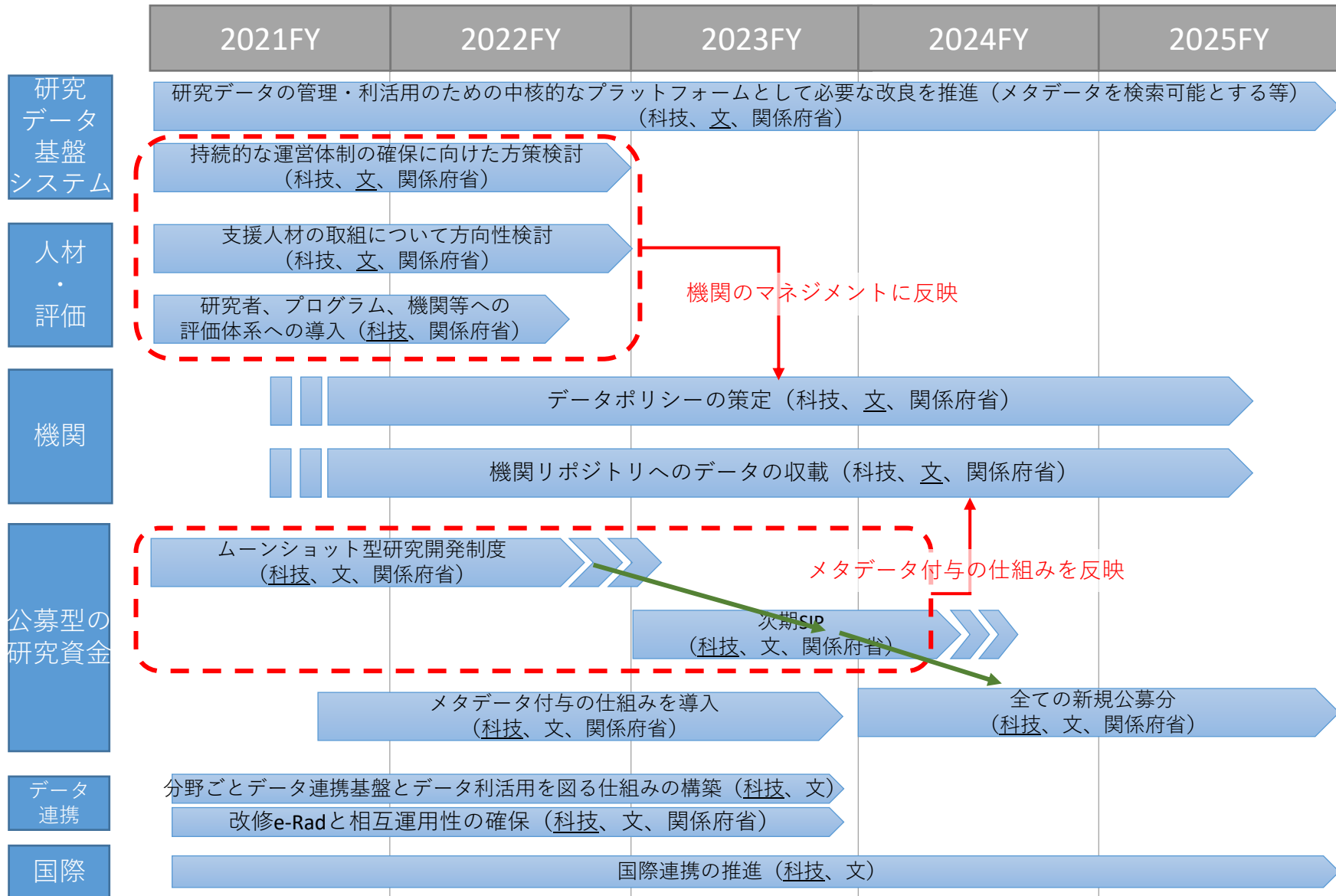
※1：国立大学法人、大学共同利用機関法人及び国立研究開発法人は、基本的な考え方において、2025年までにデータポリシーを策定することとされている。なお、国立研究開発法人については、資金配分機関である日本医療研究開発機構（AMED）、科学技術振興機構（JST）及び新エネルギー・産業技術総合開発機構（NEDO）を除く。

※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

「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」に示されたマイルストーン



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

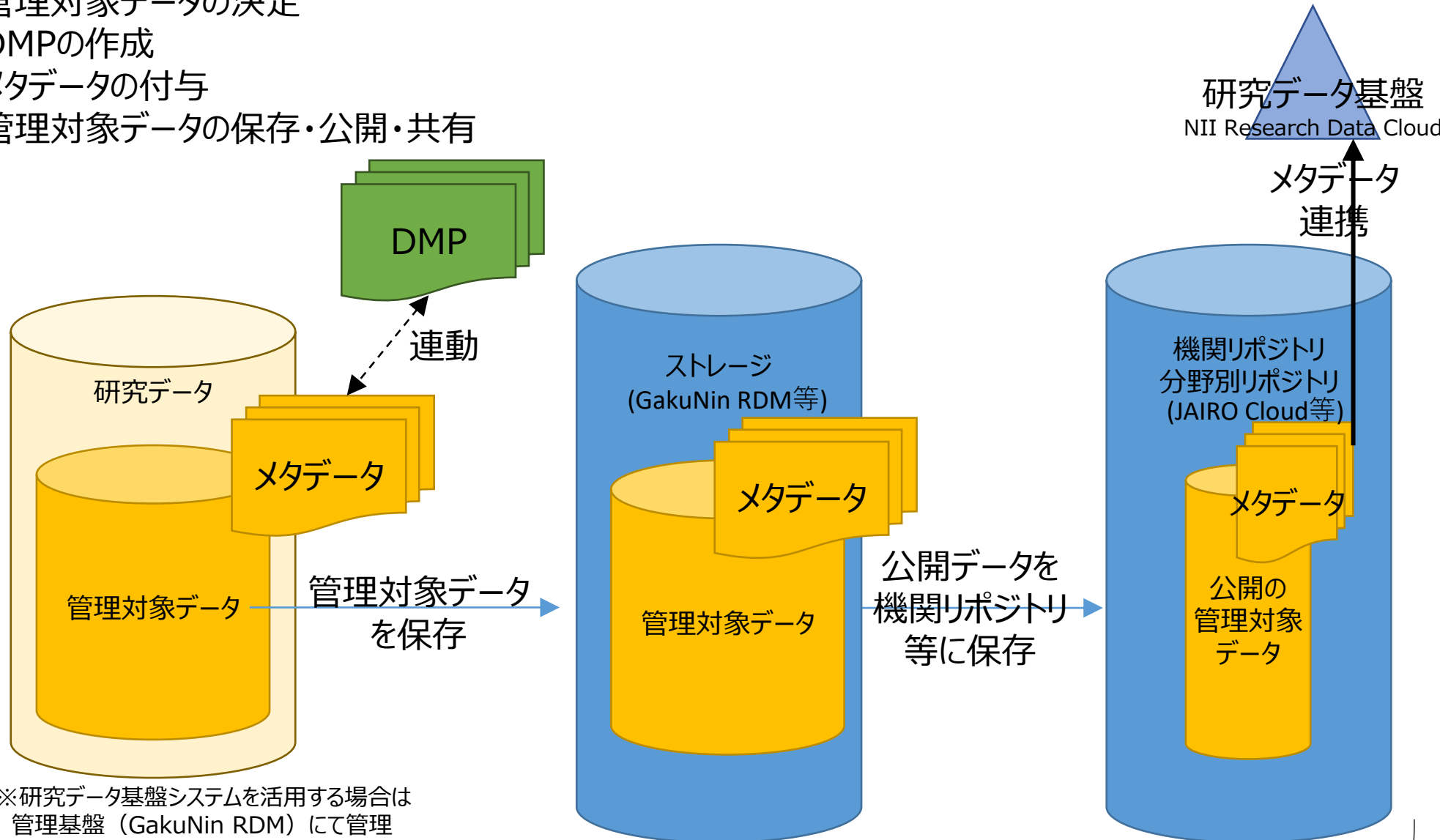
中核的なプラットフォームである研究データ基盤システム（NII-RDC）においてメタデータを検索可能な体制を構築

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、別の手段でも可。

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



(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

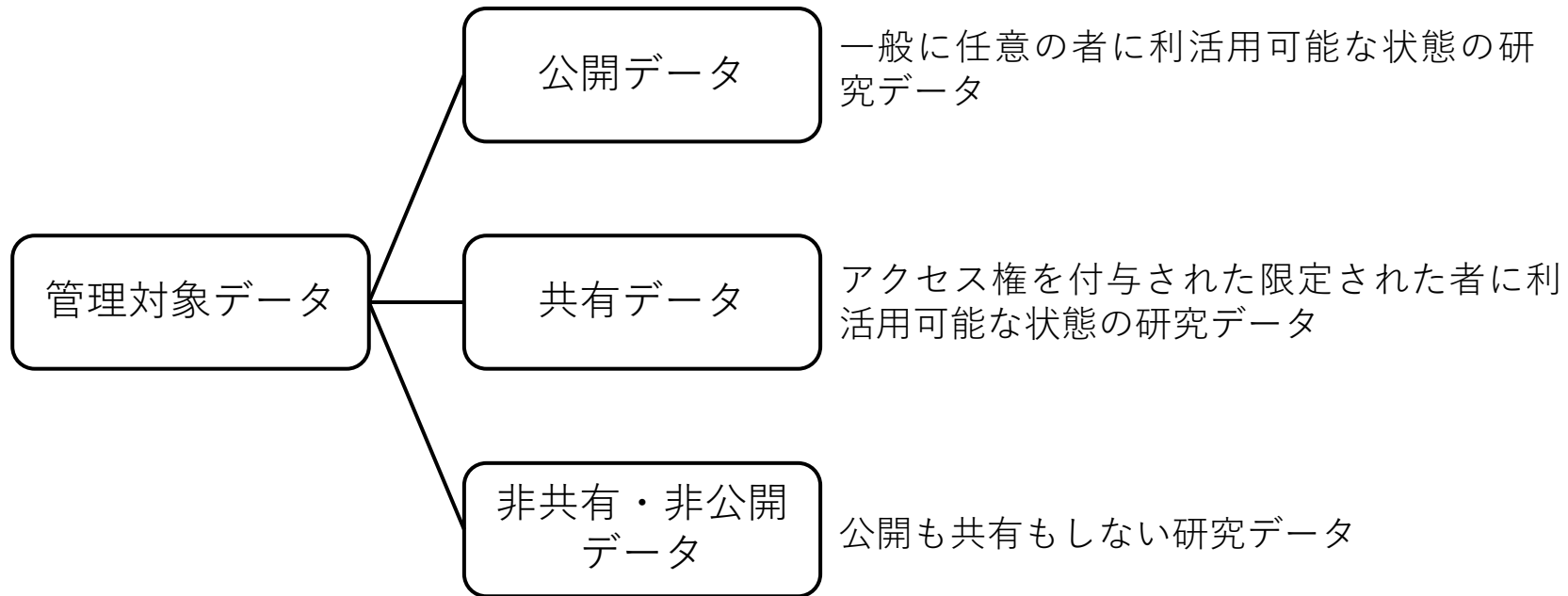
研究データのうち、研究者の所属する研究開発を行う機関や資金配分機関の基準等に基づいて、管理・利活用の対象として、研究者がその範囲を定めるものをいう。

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

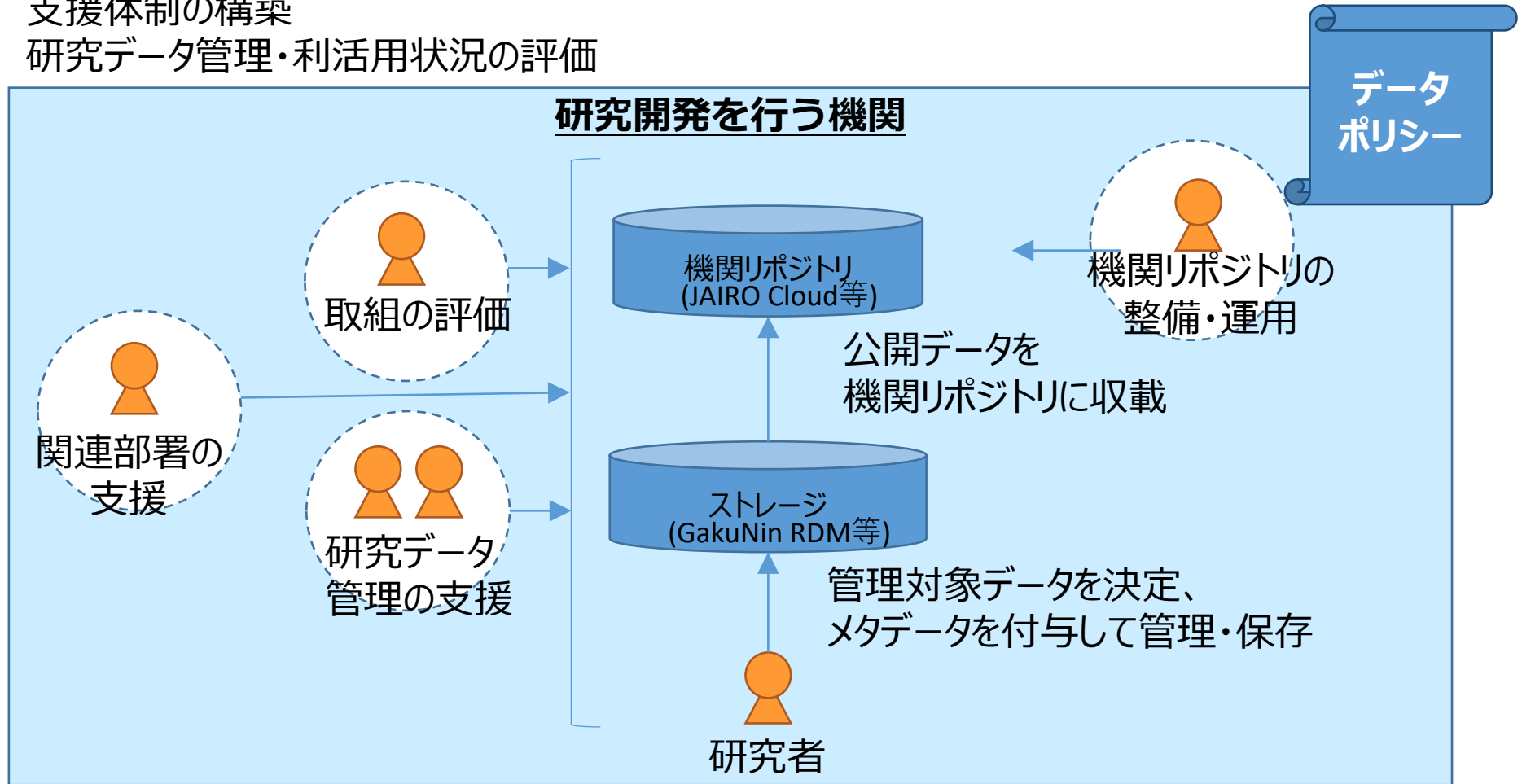


※) 「公的資金による研究データに関する基本的な考え方」から要約

- ✓ 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

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



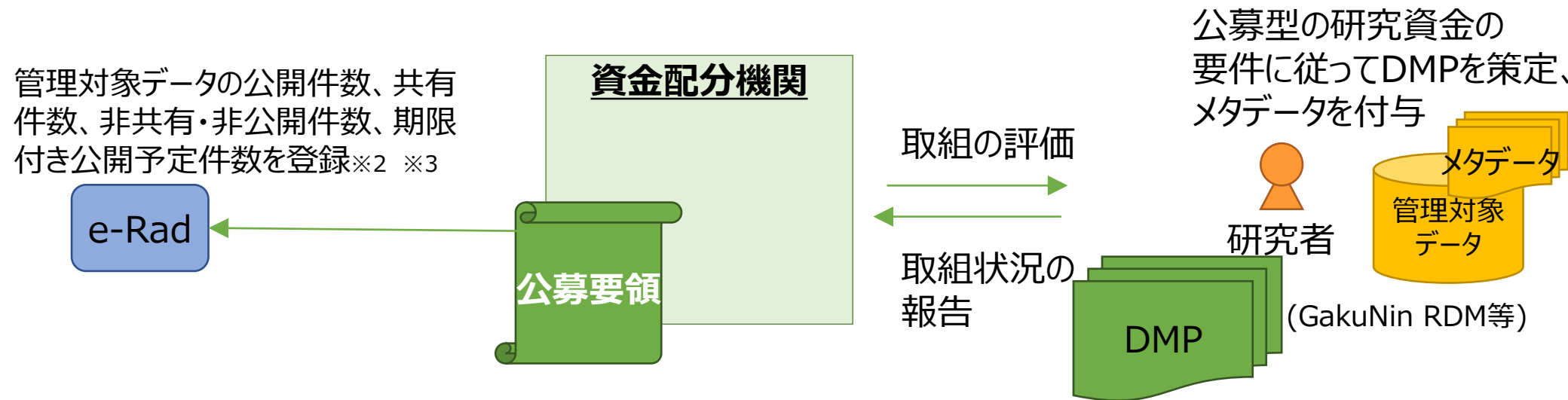
※データポリシーは機関としての研究データの管理・利活用の方針を示すもの。

※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）。

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）



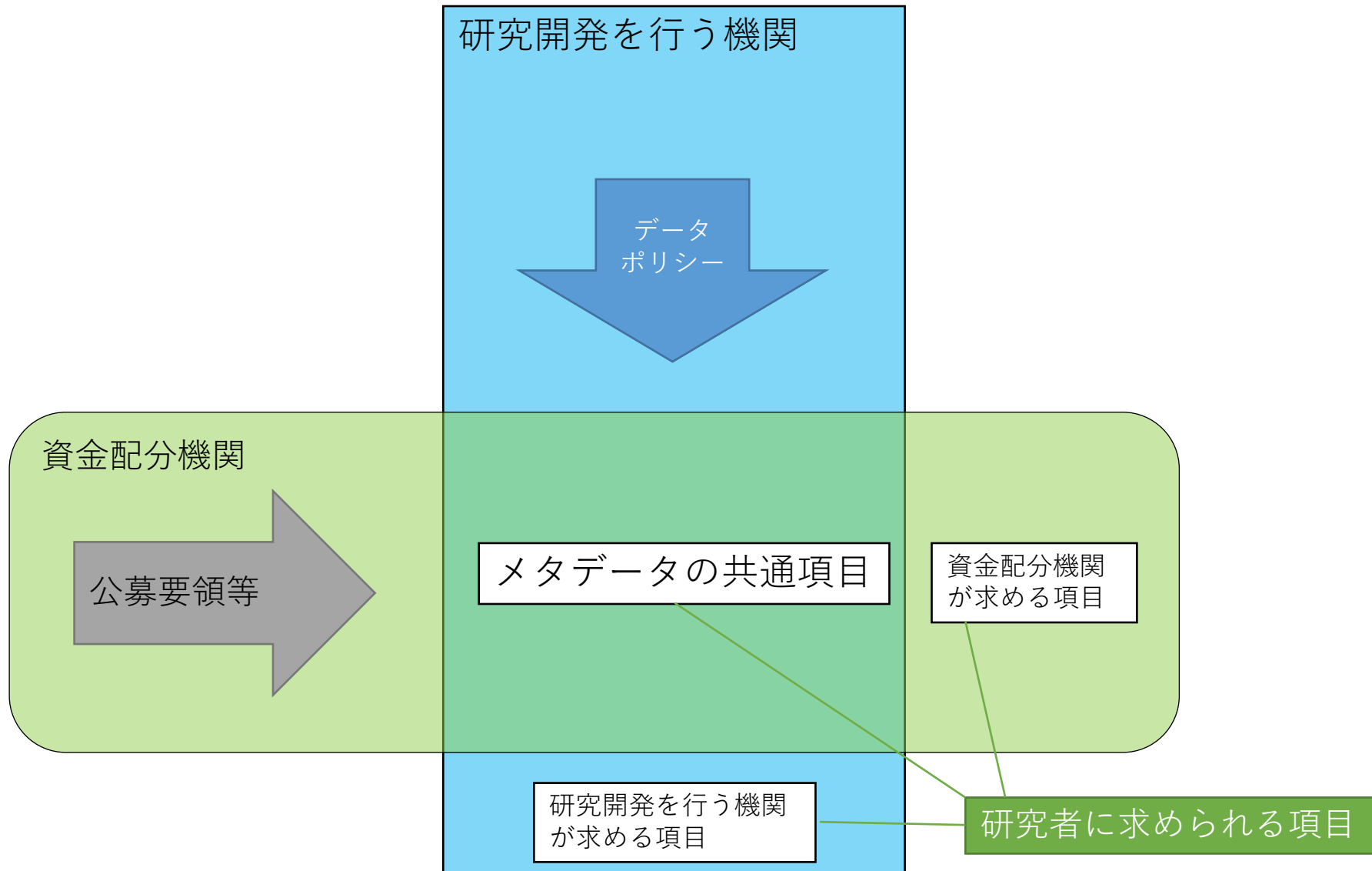
※公開する管理対象データの収載場所については、機関リポジトリJAIRO Cloud（研究データ基盤システムの公開基盤）の利用、または、研究データ基盤システムの検索基盤CiNii Researchと連携した分野別リポジトリ・汎用リポジトリへのデータの収載を推奨する。なお、これによっても対応が難しい場合には、各機関のシステムの活用等により対応する。（これらと研究データ基盤システムとの連携は別途進める）

※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。

(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



実施状況（統合イノベーション戦略2023）

【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：16機関 ※R4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：3/4 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%）
 - 2022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66%（134件のうち88制度（一部導入済み 58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：844 個（2022 年度）
- 研究データ公開の経験のある研究者割合：44.7%（2020 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2023

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年度までに体系的なメタデータ¹⁵²の付与を進め、同年度以降、研究データ基盤システム上で</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2022年度に開始し、各分野・機関の研究データをつなぐ全国的な研究データ基盤の高度化を推進し、大学等研究機関・研究者に対する研究データ基盤の利活用に向けた普及・広報活動を精力的に実施。 ・研究データ基盤の中長期的な視点での持続的・安定的運用体制の確保に向けて、将来構想を見据えた受益者負担 	<ul style="list-style-type: none"> ・引き続き、全国的な研究データ基盤（NII Research Data Cloud）の構築・高度化・実装と、AI解析等の研究データ基盤の構築・活用に資する環境の整備を行う研究DXの中核機関群を支援するとともに、2023年度から大学における研究データマネジメントにかかる体制・ルール整備の支援を実施。【科技、文、関係府省】 ・研究データ基盤の利用機関による利用状況の分析を実施し、関係機関間と
<p>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。【科技、文、関係府省】</p>	<p>モデルの在り方について検討を開始。</p> <ul style="list-style-type: none"> ・メタデータをEBPMに活用するため、e-Radでの実績報告時にメタデータの件数を登録する等の改修の検討を実施。 ・「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、概要説明や事例等を記載した資料「研究データ2022」を策定し、関係府省への周知や事例の共有等を実施。 	<p>の対話を通じた丁寧な議論により、受益者負担モデルの在り方を定め、適切な利用料金及び徴収時期等の設定検討を推進。【科技、文、関係府省】</p> <ul style="list-style-type: none"> ・e-Radの改修に合わせ、e-Radでの実績報告時にメタデータの件数を登録する等の機能改修に着手。【科技】 ・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2022」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。【科技】
<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。【科技、文、関係府省】</p>	<ul style="list-style-type: none"> ・大学等の研究開発を行う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。 ・研究データへのメタデータ付与を加速・簡略化するため、データマネジメントプランと連動した研究データ基盤におけるガバナンス機能の開発が進展。 ・研究者が一層自由に最先端の研究に打ち込める研究環境のうち研究に専念する時間の観点から、研究設備・機 	<ul style="list-style-type: none"> ・大学等の研究開発を行う機関において、引き続き研究データポリシーの策定を推進するとともに、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推進。【科技、文、関係府省】 ・「研究時間の質・量の向上に関するガイドライン」等をもとに、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引き続き推進。【科技、文】

(参考) 統合イノベーション戦略2023

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>念する時間の観点から、研究設備・機器の共用、研究データの管理・利活用の推進、U R Aや支援職員の活用促進などを盛り込んだ「研究時間の質・量の向上に関するガイドライン」を作成するとともに、これを「地域中核・特色ある研究大学総合振興パッケージ」と連動した。</p>	
<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・利活用を図るため、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S I P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>、文、関係府省】</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先進的データマネジメントのフォローアップを実施し、メタデータの付与等における課題を確認するとともに、今後の推進に向けた整理を実施。 ・S I P第3期における先進的データマネジメントの導入を検討。 ・健康・医療データ利活用基盤協議会にて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、同意の在り方を整理。 ・健康医療データ利活用プラットフォームの一部の機能について限定したユーザーによる試験的運用を開始。 ・公募型の研究資金を所管する関係府省において、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を推進。 	<ul style="list-style-type: none"> ・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を引き続き実施し、メタデータを検索可能な体制の構築を加速するため、実施状況の検証を行うとともに、自己評価及び外部評価に反映させる。【<u>科技</u>】 ・S I P第3期における先進的データマネジメントの導入について、メタデータの管理方法やシステムの連携方法等の具体化を引き続き実施。【<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(データを持ち込み扱える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を、2023年度に整備。 【<u>健康医療</u>、文、厚、経】 ・公募型の研究資金におけるDMP及びこれと連動したメタデータの付与を行う仕組みを2023年度中に導入。なお、この際には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から得られた知見を横展開する。【<u>科技</u>、文、関係府省】

(参考) 統合イノベーション戦略2023

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"> 研究データ基盤システムと分野間データ連携基盤におけるカタログサイトとの連携によるメタデータの流通を実現。 「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2022年度に開始し、各分野・機関の研究データをつなぐ全国的な研究データ基盤及び当該基盤の活用に係る環境の整備を推進。 	<ul style="list-style-type: none"> 研究データ基盤システムと分野間データ連携基盤におけるカタログサイトとの連携によるメタデータの流通環境を活用して、分野間の連携活動を推進。【<u>科技</u>】 「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、各分野・機関の研究データをつなぐ全国的な研究データ基盤及び当該基盤の活用に係る更なる環境の整備を推進する。【<u>文</u>】
<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、研究の第一線から選いたシニア人材、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>、<u>関係府省</u>】</p>	<ul style="list-style-type: none"> 科学技術・学術審議会情報委員会下に設置した「オープンサイエンス時代における大学図書館の在り方検討部会」において、大学図書館機能のデジタル化を前提とした「デジタル・ライブラリー」の実現に向け、それぞれの大学図書館が検討すべき取組の方向性を示した審議のまとめを策定。 「研究時間の質・量の向上に関するガイドライン」において、研究データの管理・利活用について行動変容を促すための観点を検討。 	<ul style="list-style-type: none"> 「オープンサイエンス時代における大学図書館の在り方検討部会」における審議のまとめで示された方向性を踏まえ、「デジタル・ライブラリー」の実現に向けて、大学図書館が今後行うべき具体的な取組について検討する会議体を設立。2030年度を目途とし、その実現に向けたロードマップを作成。【<u>科技</u>、<u>文</u>、<u>関係府省</u>】 大学等に対する支援策との連携等を通じ、研究データの管理・利活用を促進するための支援体制の整備を推進。【<u>科技</u>、<u>文</u>、<u>関係府省</u>】
<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"> 我が国がG7議長国を務める2023年のG7科学技術大臣会合に向け、G7各国における基盤システムとの相互運用性の検討をはじめ、オープンサイエンス推進のためのアセスメント・評価・インセンティブの在り方の検討を実施。 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の実現に向けて、パイロットプロジェクトにおけるシステム連携の検討を実施し、2023年5月にデモンストレーションを実施。 	<ul style="list-style-type: none"> G7科学技術大臣会合を踏まえ、G7各国におけるオープンサイエンス推進に向けた検討を引き続き推進。【<u>科技</u>、<u>文</u>】 EOSCと研究データ基盤システム間での相互運用性の推進を引き続き検討するとともに、米国等、他国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>、<u>関係府省</u>】</p>	<ul style="list-style-type: none"> 研究データの管理・利活用に関する取組の状況を評価体系へ導入するための指針として、「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について」を2022年11月に策定。 	<ul style="list-style-type: none"> 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>、<u>関係府省</u>】

研究データ管理・利活用の事例

ムーンショット型研究開発制度

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、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：265件/2週間 (8月末～9月上旬に集計)

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

アクセス数：6件

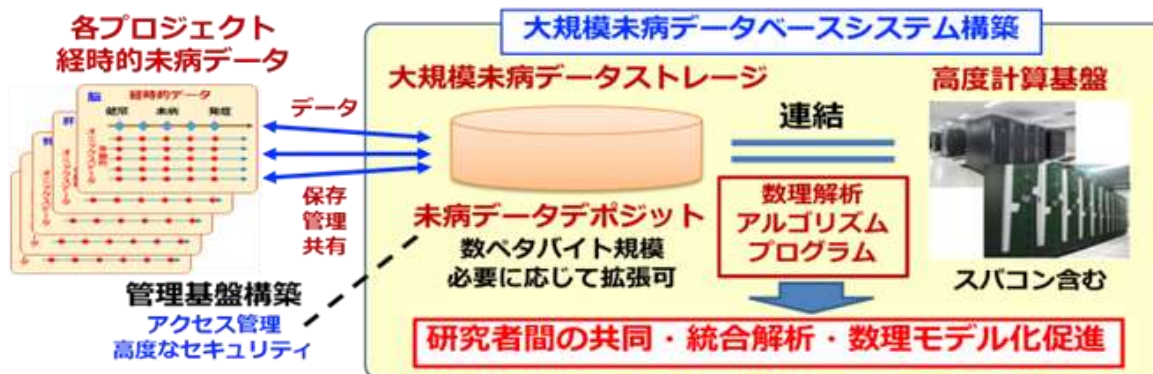
いずれも公開からの期間が短いため、具体的な活用事例についての情報は、確認できていない。日本語日常対話コーパスについては、100件弱のcloneがあると推定されており、水面下で活用されている可能性が高い。

出典：JST提供

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。

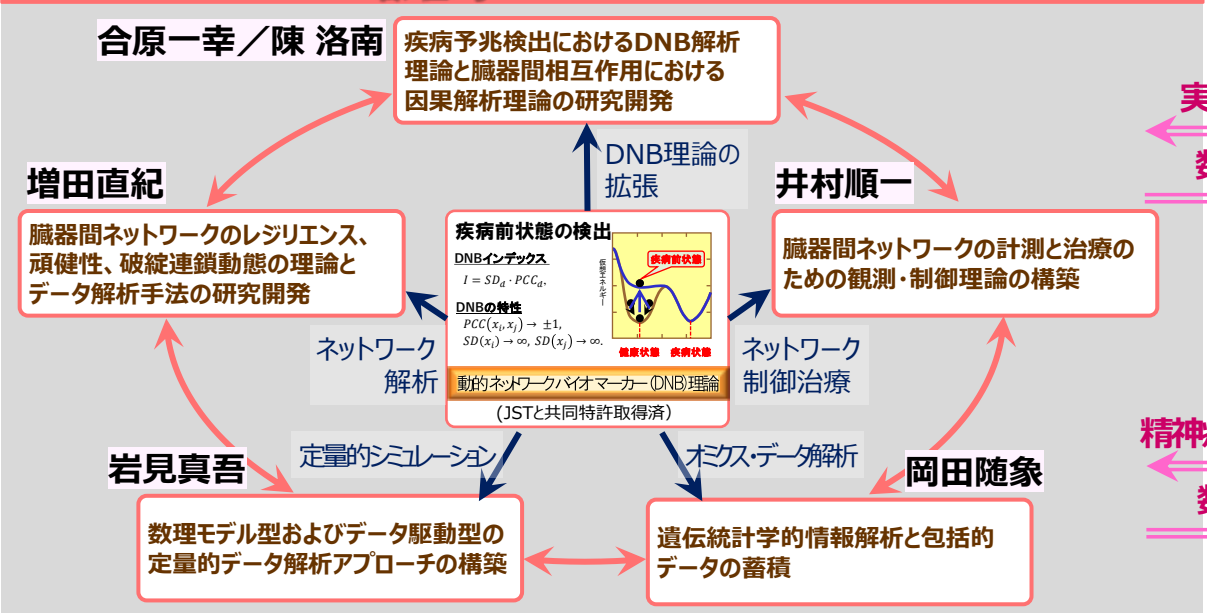
出典：JST提供

事例紹介 (JST・目標2)

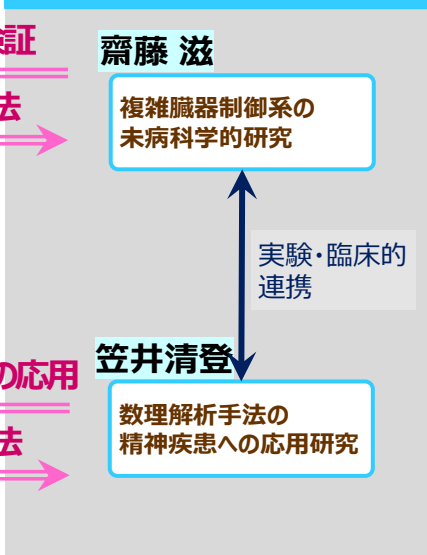
合原一幸PM 「複雑臓器制御系の数理的包括理解と超早期精密早期精密医療への挑戦」



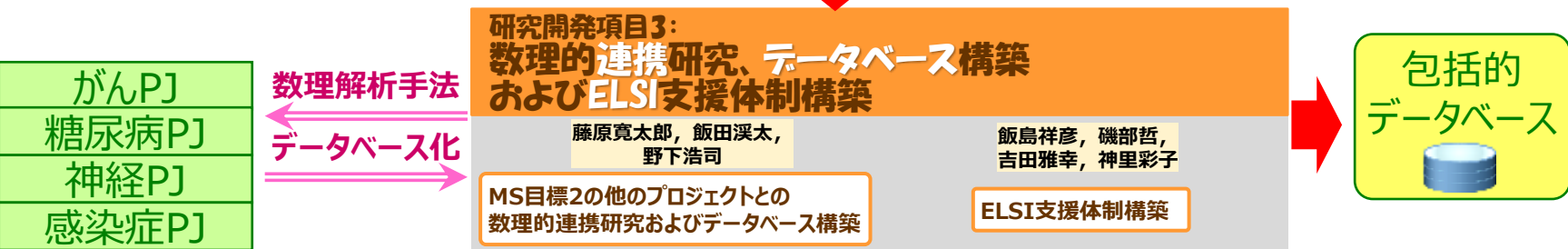
研究開発項目1: 複雑臓器制御系への数理的アプローチ



研究開発項目2: 複雑臓器制御系への実験的アプローチ



数理手法 と データ

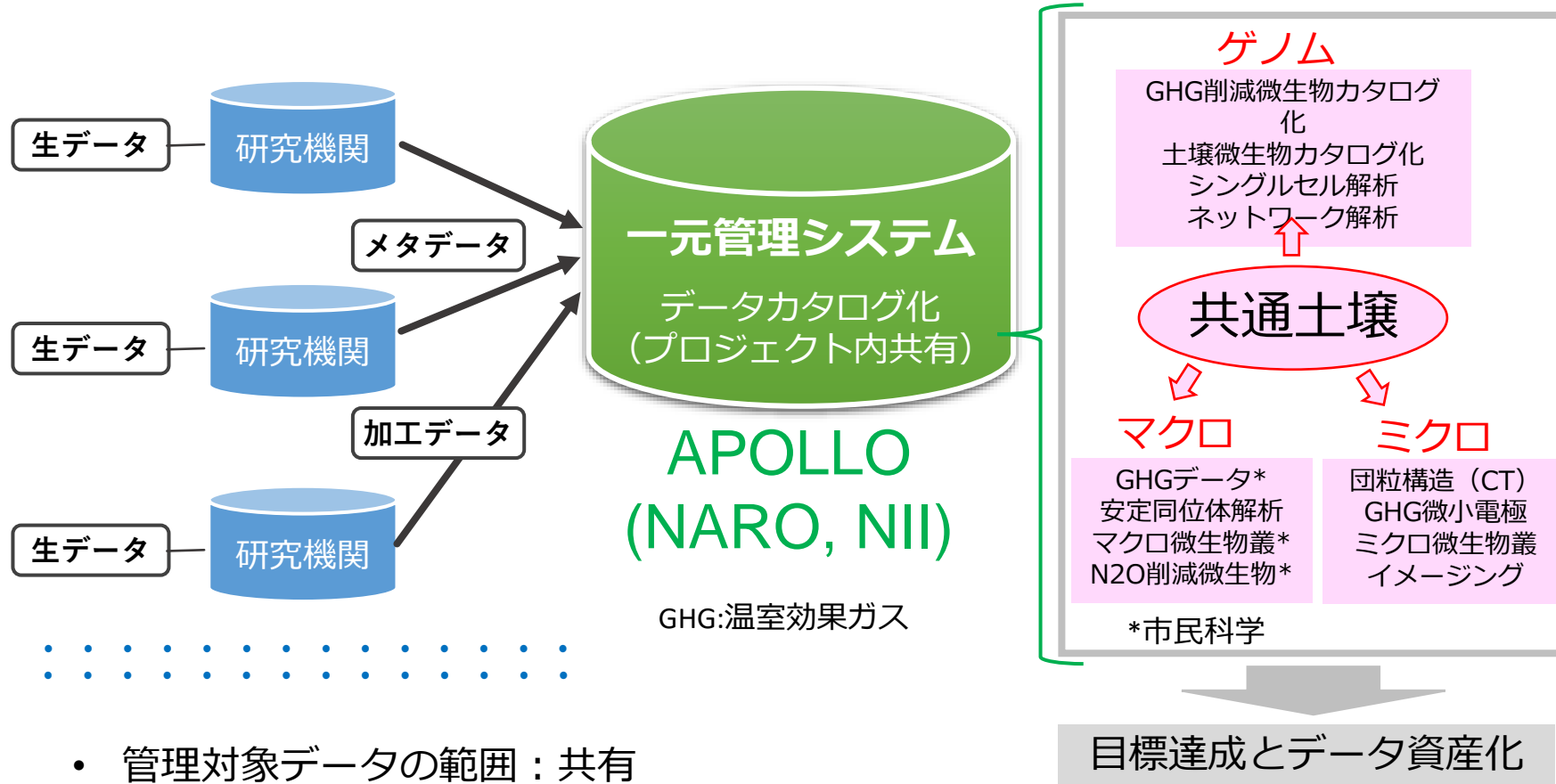


事例紹介 (NEDO・目標4)

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

事例紹介 (BRAIN・目標5)

竹山PM 「土壤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。
- 「**土壤微生物アトラス**」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。

<土壤微生物アトラス>

土壤微生物の特徴(系統・遺伝子)を土壤の性質ごとに体系化し、土壤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
- 土壤・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

- 土壤健康度の指標
- 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>

土壤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

土壤分析関連機能

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、土壤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

【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。

事例紹介 (BRAIN・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

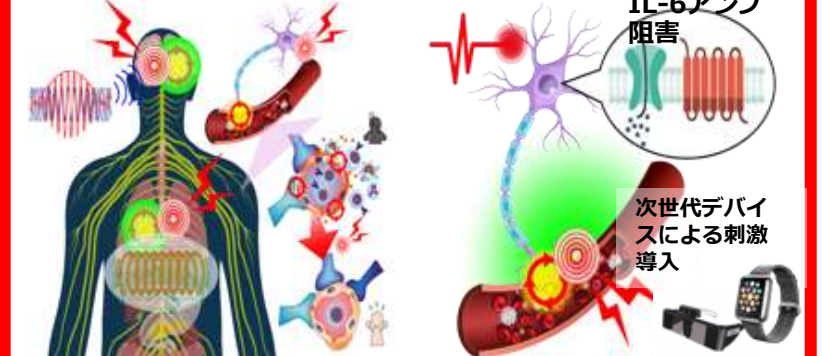
最先端の量子技術

次世代デバイスによる
検出・解析と情報発信



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



- 管理対象データの範囲
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

研究データ管理・利活用

関係府省における事例

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

新型コロナウイルス感染症の猛威により、我が国のデジタル化への遅れが顕著になったことから、次の成長の原動力として「デジタル」「AI」が最重要視されている。そのような中、AI・データ駆動型研究の重要性が高まるなど、研究手法が大きく変化しており、研究DXにより生産性を飛躍的に向上させるためには、膨大な量の高品質なデータの利活用を推進していくことが鍵である。このため、我が国における研究データの管理・利活用を促進するための中核的な研究データ基盤の構築・高度化・実装を行い、各分野等で構築が進められているデータプラットフォーム等と連携した、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・利活用を促進し、AI等の活用を推進することが求められている。

【経済財政運営と改革の基本方針2021】（令和3年6月18日閣議決定）

研究の生産性を高めるため、研究DXを推進するとともに、研究を支える専門職人材の配置を促進する。

【成長戦略フォローアップ2021】（令和3年6月18日閣議決定）

・研究のDXの実現に向け、AI・データ駆動型研究を推進するため、全国の先端共用設備や大型研究施設も効果的・効率的に活用し、2022年度からマテリアル、ライフサイエンス等多様な分野の研究データを戦略的に収集・共有・活用する取組を強化する。

未解決の課題

- 各分野におけるデータプラットフォームや、各機関におけるリポジトリの構築等が進められている。これらをつなぎ、分野・機関を越えてデータを共有・利活用するための全国的な研究データ基盤の実装が未実施であり、国際的にも遅れをとっている。
- 政府全体の方針に基づき、公的資金による研究データの取扱いに当たり、研究者に求められる責務が増大（データマネジメントプランの作成、メタデータ付与等）しており、対応が必要。
- 研究データの取扱いルール等の制度の整備や普及が追いついておらず、データサイエンスに不可欠であるデータマネジメント人材も不足。
- DXによる研究手法の変革が一部にとどまっており、デジタル基盤を徹底的に活用したAI・データ駆動型研究の進展が不十分。

実施内容

事業期間：R4年度～R8年度

- 我が国の研究力の飛躍的発展を図るため、各分野・機関の研究データをつなぎ全国的な研究データ基盤の構築・高度化・実装と、AI解析等の研究データ基盤の活用に資する環境の整備を行う、研究DXの中核機関群を支援する。

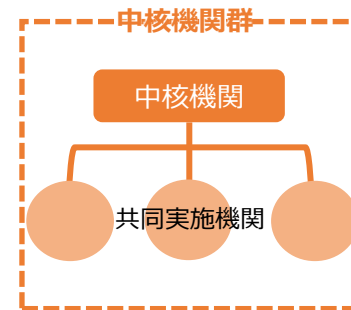
●全国的な研究データ基盤の構築・高度化・実装

- ・ユーザーニーズを踏まえながら、研究データの管理・蓄積・利活用・流通といった点で適切かつ実用的な機能を確保した全国的な研究データ基盤を整備し、AI・データ駆動型研究を推進。
- ・構築が進む各機関・各分野のリポジトリやデータプラットフォームとの連携・接続。

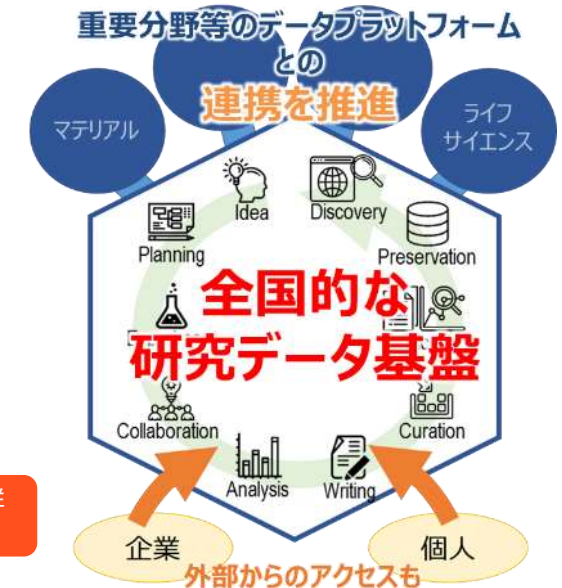
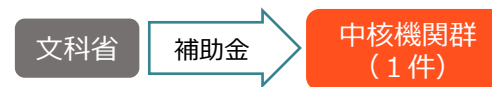
●研究データ基盤の活用に係る環境の整備

- ・効率的なAI活用のための、機械可読データの統一化や標準化等を含めたルール・ガイドライン整備、データマネジメント人材育成支援等、ユーザー視点に立って研究データ基盤を最大限に活用するための環境整備。

<実施体制>



<事業スキーム>



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

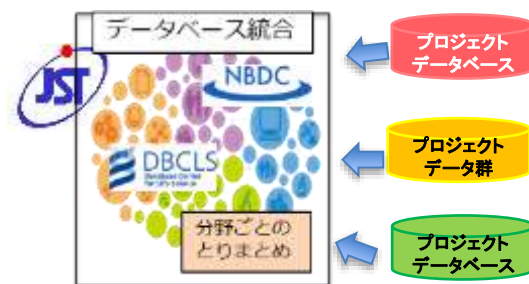
- ヒトゲノム情報や臨床情報等のデータ利活用に関する取組を推進（内閣府とりまとめ） <Visiting型計算環境、検索機能、カタログ化、ワンストップ窓口>



- 健康情報等の集約・匿名化（次世代医療基盤法（内閣府とりまとめ）活用等）

その他ライフサイエンス関係データ

- 「ライフサイエンスデータベース統合推進事業」において、効率よく研究者、産業界、さらには国民に還元していくための統合的なデータベースの構築・利活用促進
- 研究データ基盤との連携検討（データ検索機能等）



- 戦略目標「バイオDX」領域の推進とモデルとして活用
- 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

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

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

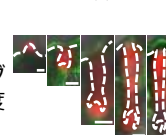
- ロボットや遠隔技術による実験の自動化
 - 計測機器・技術の高度化による高解像度・時系列のデータの大量取得
- 【取組例】
- ✓ クライオ電子顕微鏡における試料調製の自動化、操作の遠隔化
 - ✓ 理研が有する研究基盤・装置群と連携し、創薬、ワクチン関連研究や、バイオものづくりに資する分子デザイン・化学合成を推進



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）データの高效率計測・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の統合による高效率・高品質実験



事例紹介（文部科学省） 海域で発生する地震及び火山活動に関する研究開発

背景・課題

- **南海トラフにおける巨大地震**や、**突如として発生する火山噴火・火山性津波**は、我が国の国力・国勢を著しく毀損し、国民生活の水準を長期に低迷させる力を秘めた巨大災害である。
- こうした災害の対応に向けた、「**第3期海洋基本計画**」や、「**地震調査研究推進本部「第3期総合基本施策**」や、「**災害の軽減に貢献するための地震火山観測研究計画（第2次）の推進について（建議）**」等に基づき、観測システムの開発を推進していくとともに、「**地殻変動のモニタリング・予測**や**火山活動推移予測を高精度化**し、地震発生・火山噴火の長期評価の改善など防災・減災に資する成果・データを政府機関等に提供することで、被害低減に貢献していく。

（参考）各種政策文書等における防災・減災、海洋状況把握等の位置づけ

○第3期総合基本施策(R元.5)

・プレート固着・すべりの状況やスロースリップ現象に関するリアルタイムでの観測手法の開発等により、地殻変動・地震活動等の各種観測データの充実が必要である。また、物理モデルに基づく現状把握、地殻変動・地震活動データと現実的なモデルに基づいたシミュレーションを活用した（中略）予測手法の高度化が必要である。

○国土強靱化年次計画2021(R3.6)

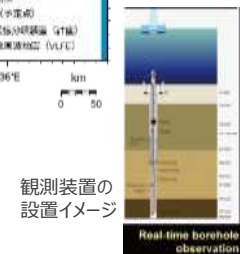
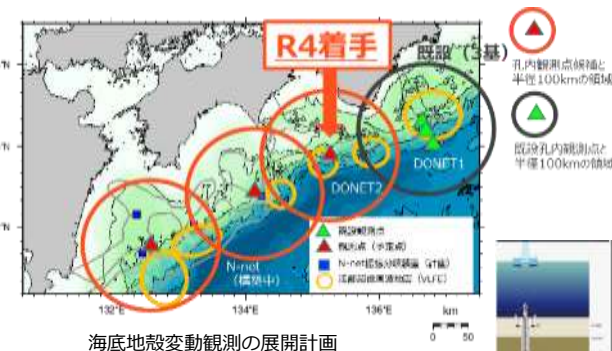
・南海トラフ西側の領域など観測網が手薄なエリアにおける観測網の整備を進めるとともに、DONETとS-netの観測データの利活用を推進する。

○災害の軽減に貢献するための地震火山観測研究計画（第2次）の推進について（建議）（H31.1）

・地震・火山現象を解明し、予測の高度化を推進するとともに、その成果を活用して地震や火山噴火による災害の軽減につながる研究を推進

事業概要

- **連続リアルタイム海底地殻変動観測技術の開発・展開** 562百万円（272百万円）
 - 南海トラフ巨大地震に向け、**地殻に蓄積されつつある歪（ひずみ）の量（地殻変動量）を広域で把握するため、海底地殻変動観測装置を開発・展開**し、発生予測の高精度化に貢献する。
 - 光ファイバーによる長期・定常的な地殻変動の推移を**高密度・リアルタイムで把握するための新たな観測技術の開発**を行う。
 - 南海トラフ西側のゆっくり滑りをリアルタイムに観測監視することを目指し、令和3年度補正予算により整備する**海底深部における地殻変動観測装置**の設置に向け**地球深部探査船「ちきゅう」による掘削**に着手する。
※令和3年度補正予算に「ゆっくり滑り（スロースリップ）」をはじめとした海底地殻変動観測装置の開発費（7億円）を計上（令和4-5年度において「ちきゅう」により掘削孔を生成し、当該観測装置を設置予定）
- **海底震源断層の高精度広域調査** 800百万円（805百万円）
 - 「かいめい」による**地下構造調査**等によって、多様な地震活動を規定する断層形状や、応力状態や滑りやすさの指標など**地下構造の実態を把握**する。
 - 地震の長期評価の更なる精度向上に不可欠な「**地震発生履歴**」を**適切に把握**するため、南海トラフ・千島海溝沿いにおいて「かいめい」によるコアリングを実施し、**長期間の地層記録により地震発生の時間分布を明らかに**することを目指す。
- **プレート固着状態・推移予測手法の開発・評価** 38百万円（38百万円）
 - 調査によって得られるより現実的な地殻構造を取り入れたモデルを構築し、より**高精度な地殻変動・津波シミュレーションを実施**するとともに、**プレート固着・すべり分布の現状把握とその推移予測手法を開発**する。
- **海域火山活動把握のための研究開発** 22百万円（22百万円）
 - 突如として発生する火山噴火・火山性津波被害の軽減に資するために、**海域火山の活動の現状と履歴を明らかに**する。また、地震・電磁気構造探査、海底試料の解析で得られた知見やデータを政府機関等に提供する。



気候変動適応戦略イニシアチブ 地球環境データ統合・解析プラットフォーム事業

背景・課題

- データ駆動型社会であるSoceity5.0では、様々なビッグデータ、リアルタイムデータは極めて重要な価値を有しており、その利活用により新しい価値の創造が見込まれる。
- 防災対策や脱炭素対策等の様々な気候変動対策において、これまでの過去データに加え、科学的な将来予測データも活用した対策へのパラダイムシフトが加速する中で、地球環境分野において、防災・減災や脱炭素社会の実現等にも貢献する地球環境ビッグデータ（観測情報・予測情報等）の利活用を更に推進することが必要。

【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）>
- ・組織を越えた防災情報の相互流通を担うSIP4Dを核とした情報共有システムの都道府県・市町村への展開を図るとともに、地域の防災力の強化に取り組むほか、データ統合・解析システム（DIAS）を活用した地球環境ビッグデータの利用による災害対応に関する様々な場面での意思決定の支援や、地理空間情報を高度に活用した取組を関係府省間で連携させる統合型G空間防災・減災システムの構築を推進する。
- <2050年カーボンニュートラルに伴うグリーン成長戦略（令和3年6月）>
- ・データ統合・解析システム（DIAS）等を通じてGHG観測データ、気候変動予測情報等の更なる利活用を推進し、科学基盤の充実を図る。

事業概要

【事業の目的・目標】

- 地球環境データを蓄積・統合解析するデータ統合・解析システム（DIAS）を活用した地球環境分野のデータ利活用を推進するとともに、国、自治体、民間企業等の気候変動対策を中心とした意思決定に貢献する地球環境データプラットフォーム（ハブ）の実現を目指す。

【事業概要・イメージ】

- ・ 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推進。
- ・ これまでの成果を生かし、GEO（地球観測に関する政府間会合）やIPCC（気候変動に関する政府間パネル）等を通じた国際貢献、学術研究を一層推進。
- ・ 地球環境ビッグデータの利活用を強化するための計算資源等の設備整備や利用拡大等を推進。



【事業スキーム】

- ✓ 支援対象機関：JAMSTEC
- ✓ 事業期間：令和3年度～令和12年度



【これまでの成果】 ※前身事業の成果

- ✓ 国内外の幅広い利用者・分野での活用により、気候変動研究等の取組が加速。
※ユーザー数が直近5年で約5倍
- ✓ 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や台風等による洪水予測をDIAS上で解析。
- ✓ DIASに蓄積されている気候変動予測データ、マラリア患者数データ等を統合解析し、マラリア流行のリアルタイム予測を実施



事例紹介（文部科学省） 気候変動対策・レジリエンスに貢献するデータ解析プラットフォームの形成

【概要】

気候変動やそれに伴う極端気象の激甚化・広域化、及び地震・津波・火山等の自然災害への対応が必要であるが、研究開発・意思決定・ファイナンスに必要なインテリジェンスの不足が大きな課題となっている。

レジリエントで安全・安心な社会やカーボンニュートラルの実現に向け、**研究開発・実務現場での地球環境ビッグデータ利活用の促進やインテリジェンスの創出**を目指し、新たな技術・価値を生み出すデータ・解析プラットフォームを形成する。



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



GERDaMSとGEDの管理
データベース共通化

データベース

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