

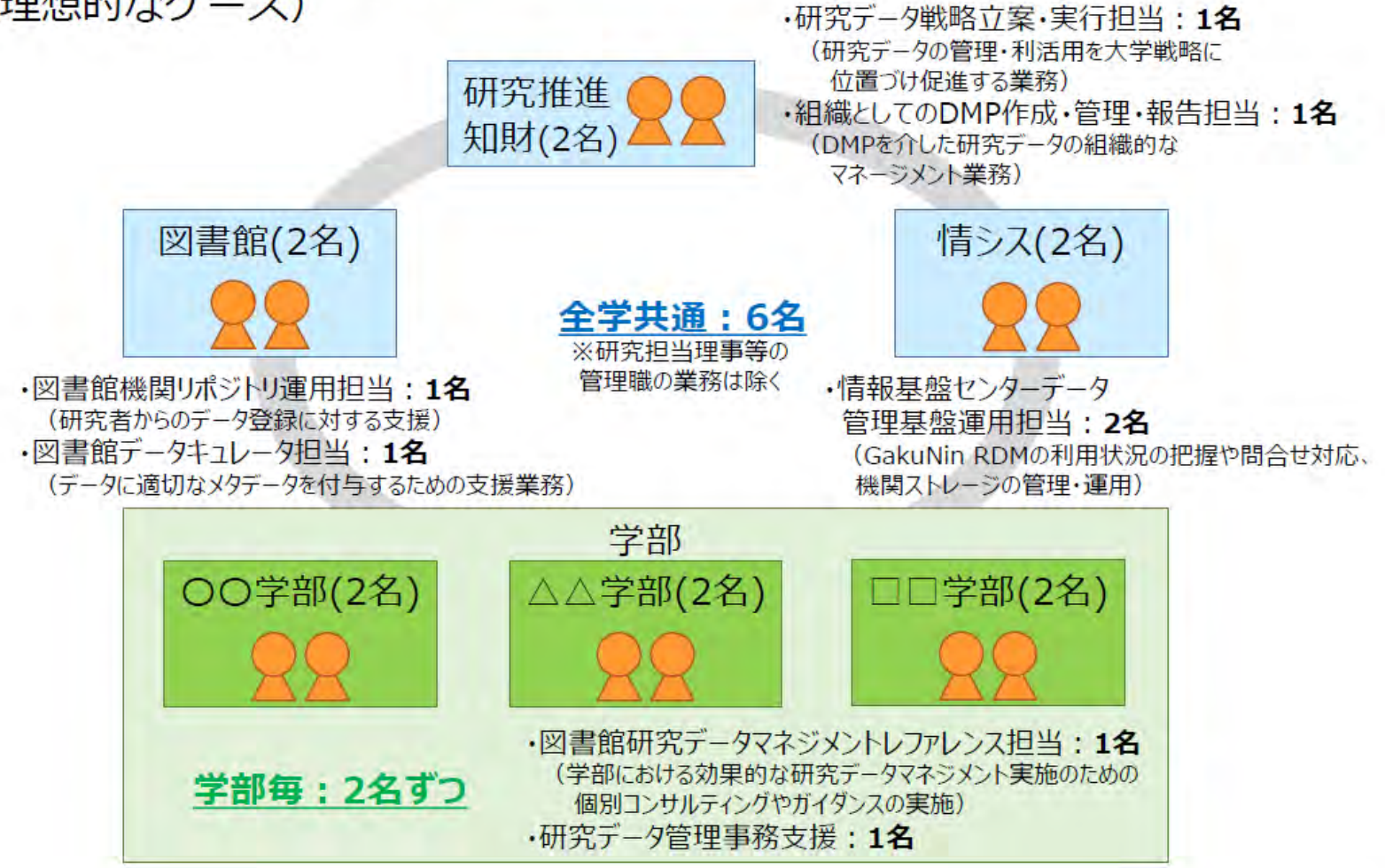


# 研究データマネジメント業務の規模感（イメージ）

中間まとめ

ガイドライン

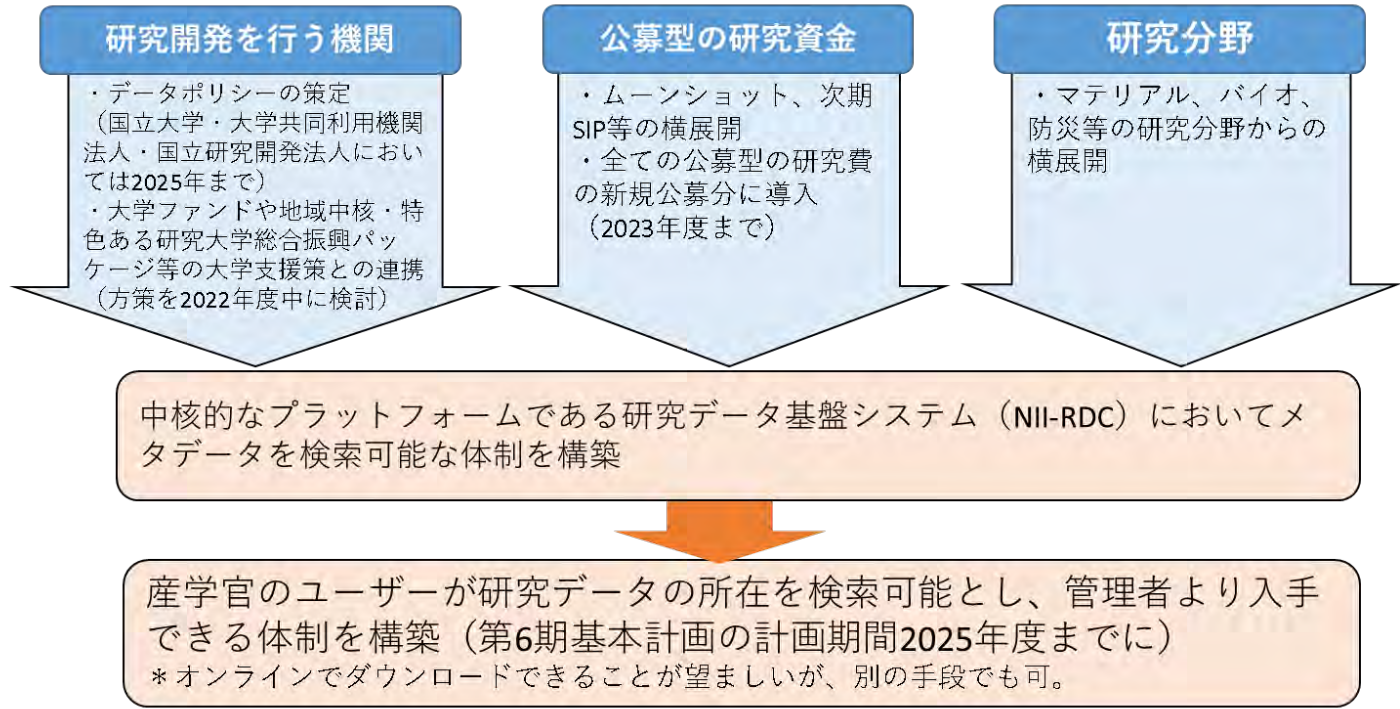
5,000人程度の大学の場合  
(理想的なケース)





## 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



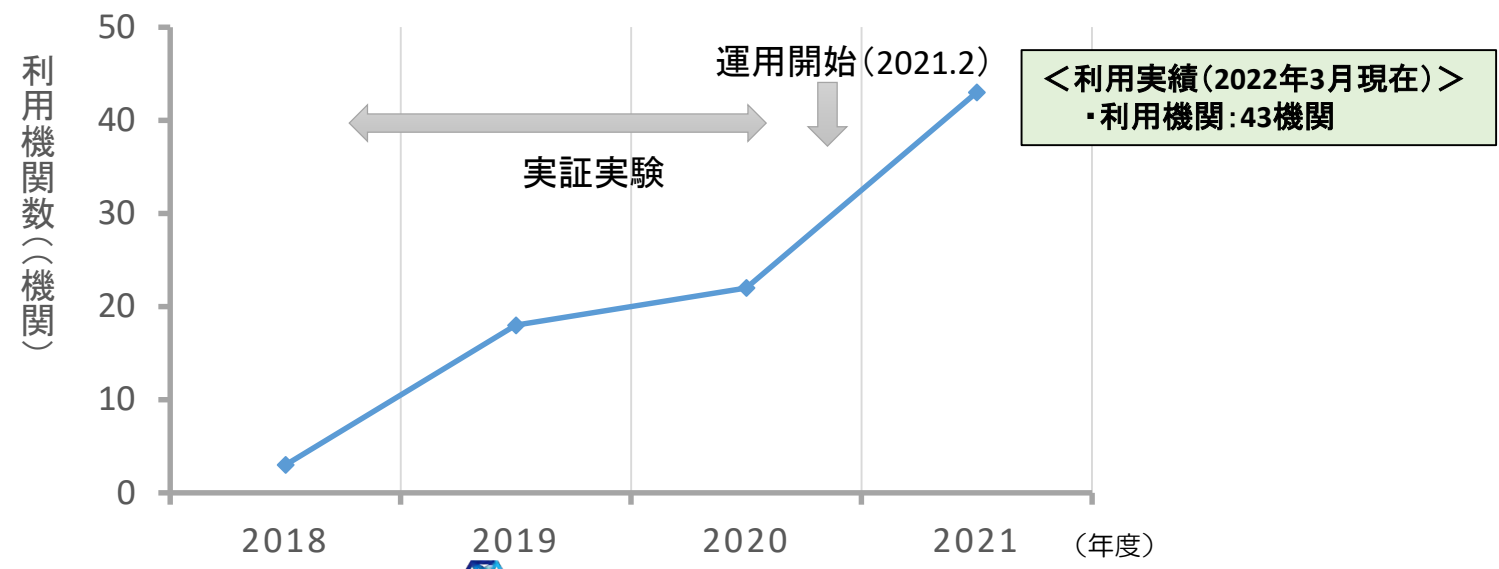


# 研究データ基盤システムの利用実績(1)

中間まとめ

**GakuNin RDM (Research Data Management)**  
 国立情報学研究所が構築し、2020年度(2021年2月本格運用開始)より運用・公開を開始した新サービス。研究プロジェクト実施中に、個人の研究者あるいは研究グループが研究データや関連の資料を管理するための研究データ管理基盤であり、既存のストレージや研究ソフトウェアと連携し、クローズドな空間で、研究プロジェクトに関わるファイルのバージョン管理や、メンバー内でのアクセスコントロールが可能。

<GakuNin RDMの利用状況 (利用機関数) >



内閣府 科学技術政策担当大臣等政務三役と総合科学技術・イノベーション会議有識者議員との会合(令和4年9月1日) 研究に専念する時間の確保について(中間まとめ案)より抜粋





## 研究データ基盤システムの利用実績(2)

中間まとめ

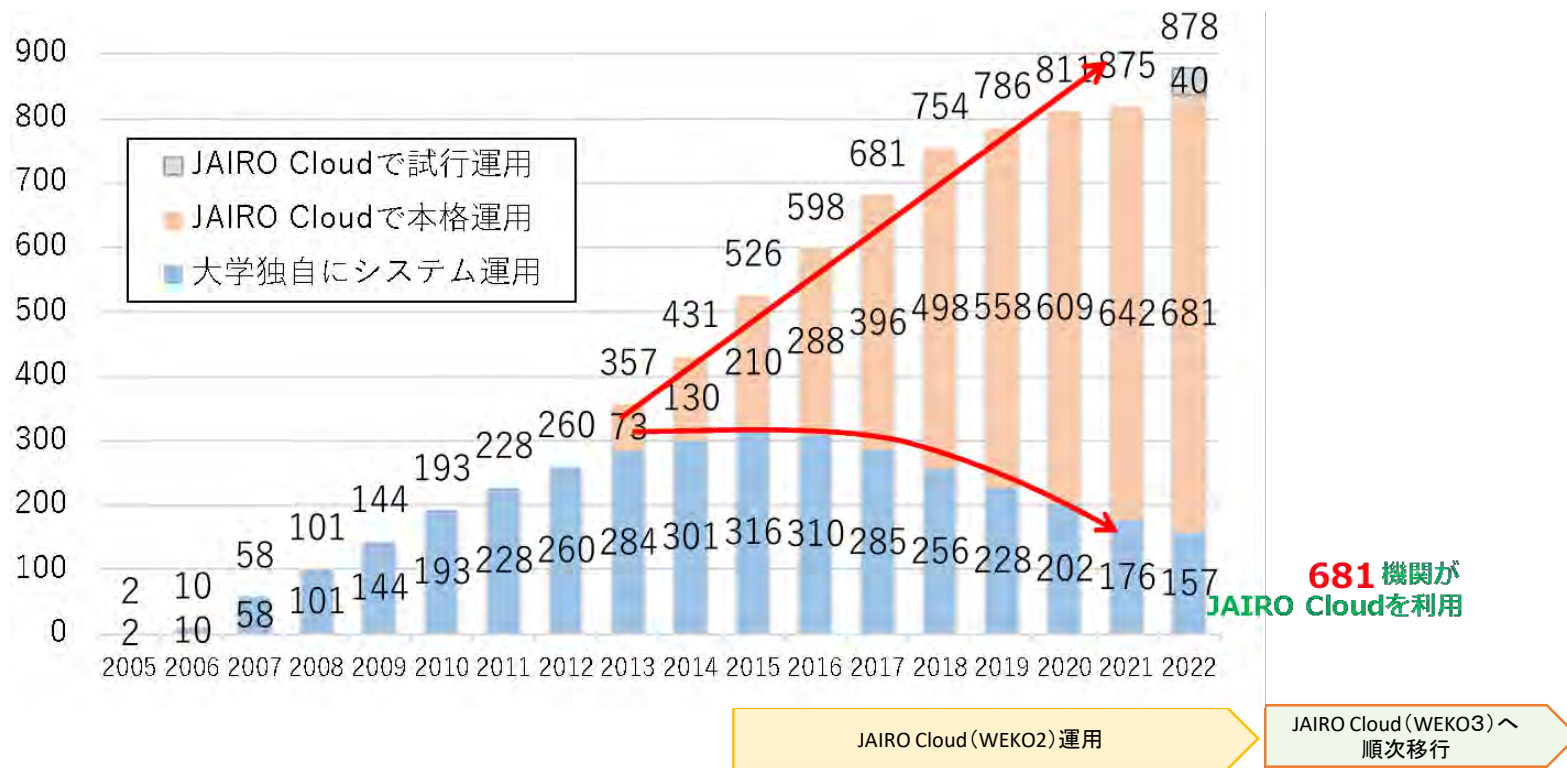
### JAIRO Cloud

国立情報学研究所が構築し、2012年より運用・公開を開始したクラウド型の機関リポジトリ環境を提供するサービス。

#### <利用機関数推移>

- 2012年度の73機関から、1年後の2013年度には、130機関に増大。その後、コンスタントに利用機関は増大し、近年では、独自構築からJAIRO Cloudに移行する機関も増大傾向
- 研究データ基盤システムとして、2020年度にWEKO3(アプリケーション)を開発※し、バージョンアップ。2022年度現在、利用機関はWEKO2からWEKO3へ順次移行手続中。

※WEKO3では、バージョン管理機能や研究データに対応したメタデータスキーマであるJPCOARスキーマへ対応するなど研究データをコントロールしやすい仕様へとバージョンアップ





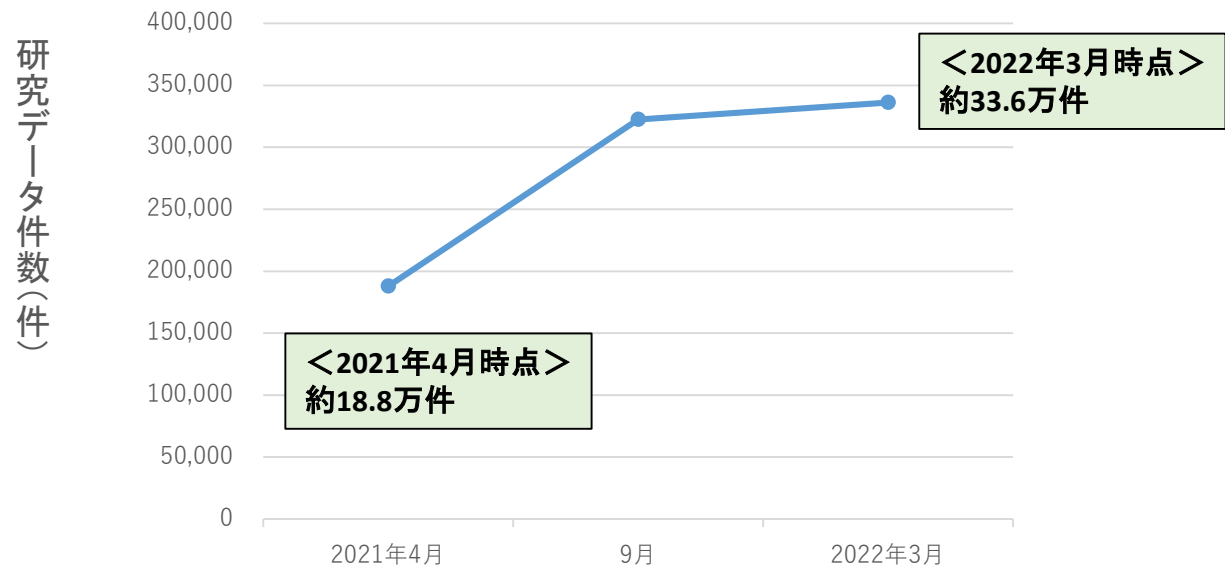
## 研究データ基盤システムの利用実績(3)

中間まとめ

### CiNii Research

国立情報学研究所が構築し、2021年より運用を開始した学術情報検索サービス。公開基盤(JAIRO-Cloud)等に登録された研究成果や論文情報のみならず、図書、研究データ、それらの成果を生み出した研究者、研究プロジェクトの情報などを包括して探索することが可能。

### <CiNii Research で検索可能な研究データ収載数>

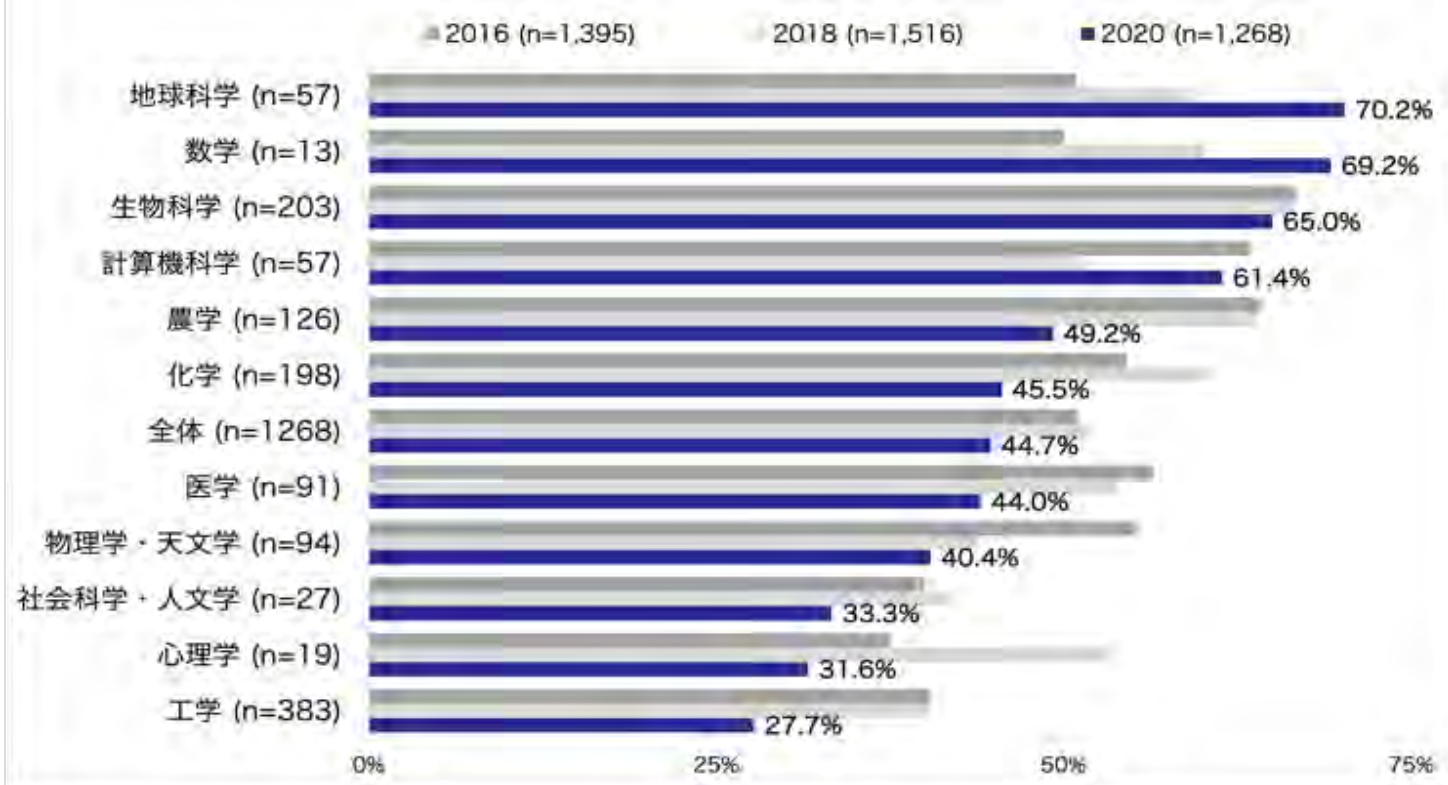


※CiNii Researchでは、これまでNIIが整備してきたサービス(CiNii ArticlesやCiNii Books等)で収載したデータを 内部で統合して、横断検索を可能にしている。

※研究データ以外の情報(論文、図書、その他の成果物等)についても2022年3月にかけて増加している。



# 分野別データ公開経験



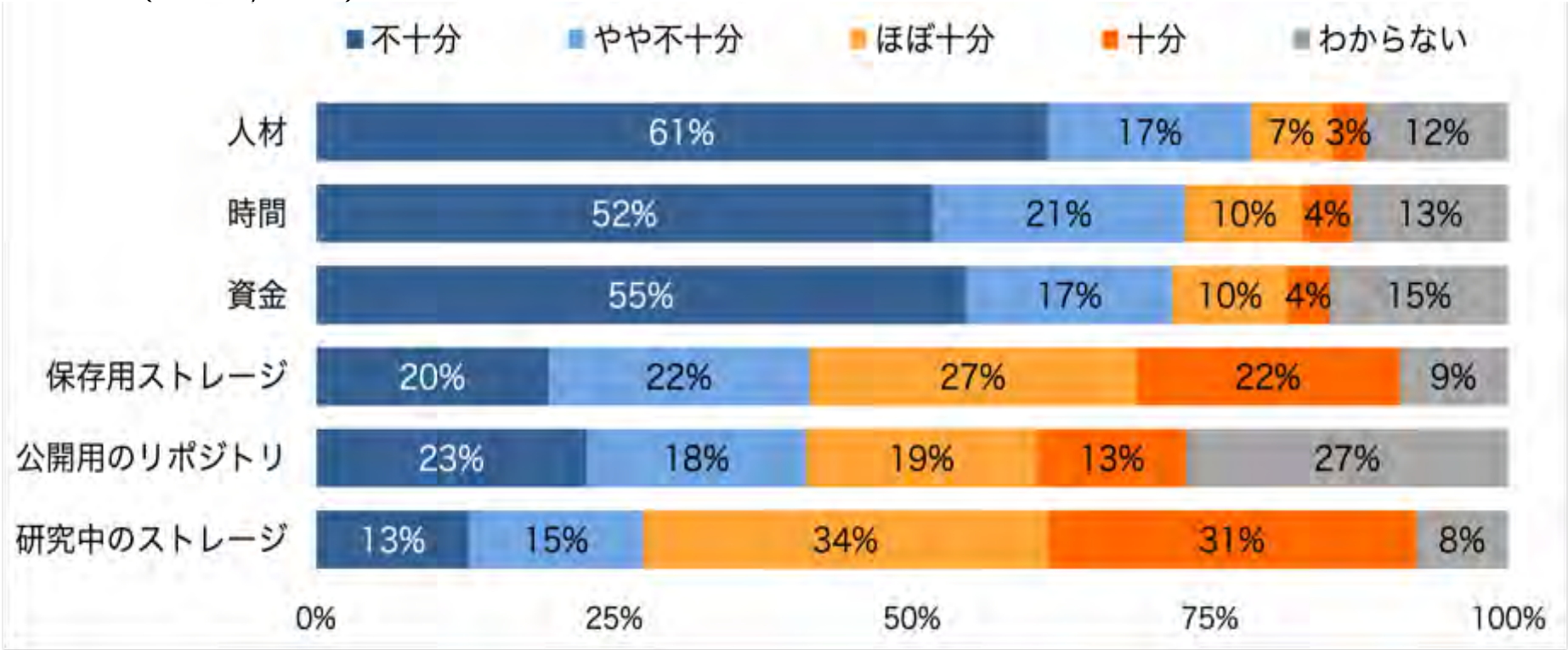
- 分野による差が大きい
- 地球科学（1位）と数学（2位）は前回調査から増加。ただし回答者数が比較的少ないため少数の回答が影響している可能性あり

出典：科学技術・学術政策研究所「研究データ公開と論文のオープンアクセスに関する実態調査2020」,  
調査資料 (Research Material) 316, NISTEP, 2021-11( <https://doi.org/10.15108/rm316> )



# データの整備・公開資源の充足度

(n=1,188)



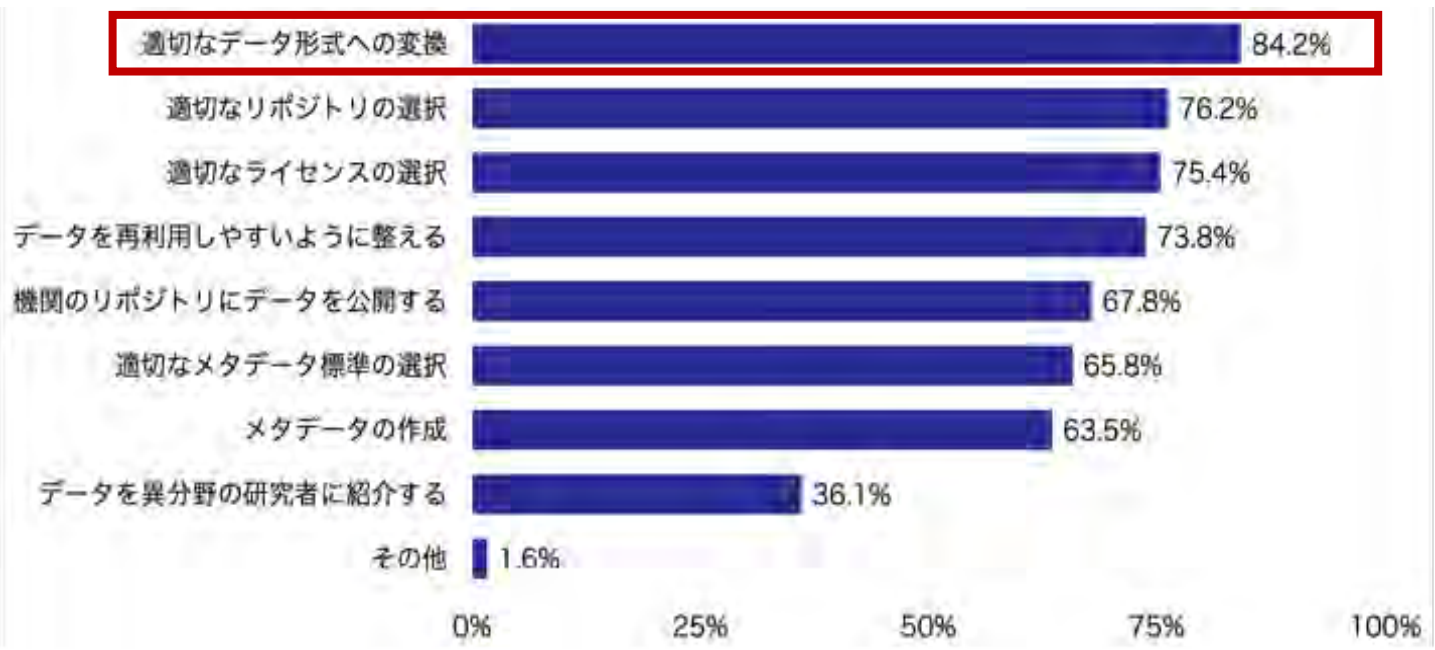
- 人材・時間・資金の不足感が強い
- ストレージやリポジトリは2016/2018年と比較して、やや改善しているものの十分とは言い難い
- 公開用のリポジトリは「わからない」の比率が高い

出典： 科学技術・学術政策研究所「研究データ公開と論文のオープンアクセスに関する実態調査2020」, 調査資料 (Research Material) 316, NISTEP, 2021-11( <https://doi.org/10.15108/rm316> )





# データの整備・公開・保存の依頼意思



- データの整備・公開・保存プロセスを図書館員やデータキュレーターに依頼したいと考える回答者は41.1% (n=1,188)
- 依頼したい項目1位は「適切なデータ形式への変換」
- 2016/2018年調査では、第三者が支援する場合に専門性が必要であると考えられる項目の1位が「適切なデータ形式への変換」であった→専門性が高いことであっても依頼したいと考えている可能性

出典： 科学技術・学術政策研究所「研究データ公開と論文のオープンアクセスに関する実態調査2020」, 調査資料 (Research Material) 316, NISTEP, 2021-11( <https://doi.org/10.15108/rm316> )





# 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」が最重要視されているが、AI・データ駆動型研究開発に必要な大規模かつ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、全国に散逸する研究データをつなぎ、必要なデータを利活用できる環境を整備することが必要。

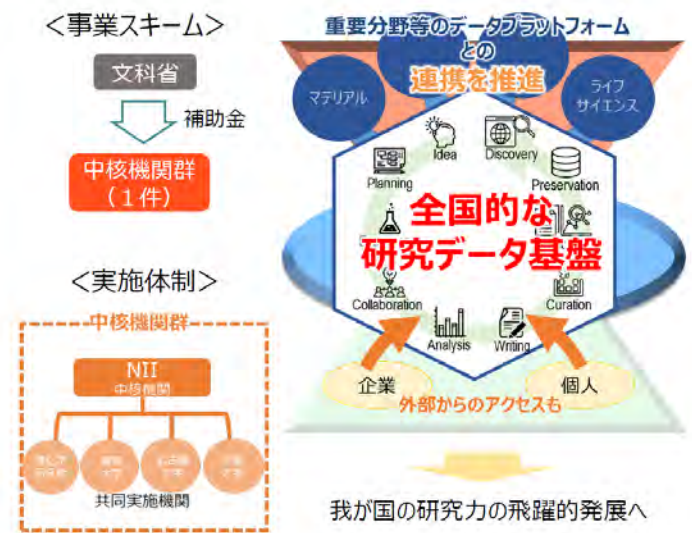
### 未解決の主な課題

- ✓ 分野・機関を越えた全国的な研究データ共有・利活用の基盤が未実装
  - ✓ 各大学のデータマネジメント体制やルール構築が進んでいない（研究データマネジメントポリシーを策定している大学は国立大学でも全体の2割程度かつ必ずしも全学的な内容になっていない）
  - ✓ DXによる研究手法の変革、AI・データ駆動型研究の進展が不十分
- 【経済財政運営と改革の基本方針2022】（令和4年6月7日閣議決定）
- ・国際性向上や人材の円滑な移動の促進、大型研究施設の官民共同の仕組み等による戦略的な整備・活用の推進、情報インフラの活用を含む研究DXの推進、各種研究開発事業における国際共同研究の推進等により、研究の質及び生産性の向上を目指す。
- 【新しい資本主義のグランドデザイン及び実行計画】（令和4年6月7日閣議決定）
- ・大学等や国の機関が保有するデータは、それぞれの機関に分散し、データ形式もバラバラとなっているが、他のデータ基盤との接続を可能とし、民間企業等の利活用を進める。

## 必要な取組 事業期間：R4年度～R8年度

①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構築・高度化・実装**と、②**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。

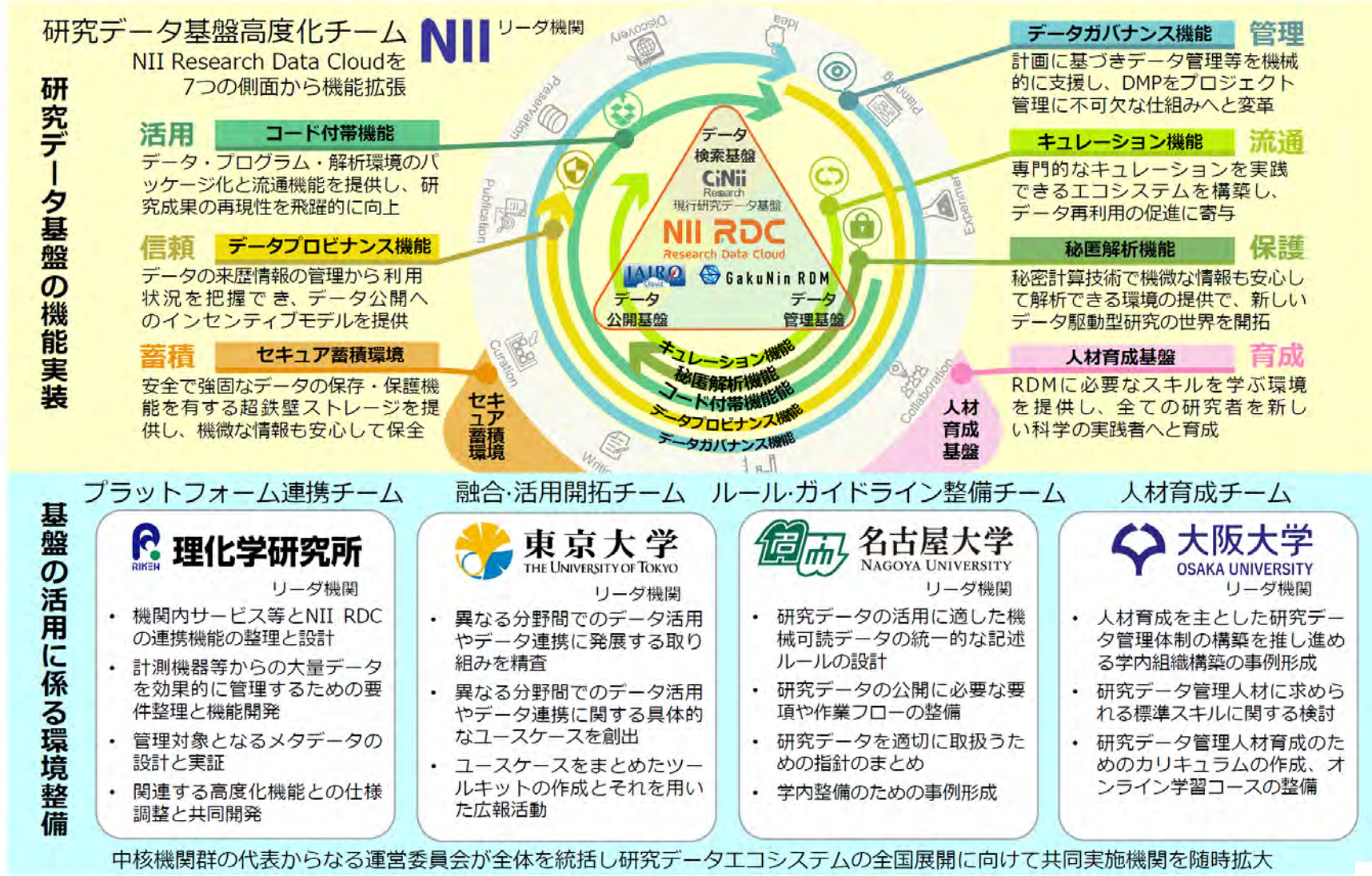
- ① 全国的研究データ基盤の構築・高度化・実装**
- 以下の機能を有した全国的な研究データ基盤の整備を推進
- ・管理データの取捨選択やメタデータ付与、データの出所・修正履歴の管理など、研究データ管理にかかる作業を効率的に実施するための機能
  - ・秘匿すべき情報が含まれる研究データを適切・安全に解析可能とする環境構築を支援する機能
  - ・各機関リポトリや各分野のデータプラットフォームとの連携・接続による分野・機関を越えた研究データの検索機能 **(拡充)** 等
- ② 研究データ基盤の構築・活用に係る環境の整備**
- ・効率的なAI活用のための、機械可読データの統一化や標準化等を含めたルール・ガイドライン整備、データマネジメント人材育成支援
  - ・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 **(新規)**
  - ・異分野等の多次元データ解析など最先端の数理・情報科学と分野の融合によるAI・データ駆動型の新たな研究開発手法の開拓 **(拡充)** 等





研究DX  
研究データの  
管理・利活用

## 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



内閣府 科学技術政策担当大臣等政務三役と総合科学技術・イノベーション会議有識者議員との会合（令和4年9月1日）  
研究に専念する時間の確保について（中間まとめ案）より抜粋



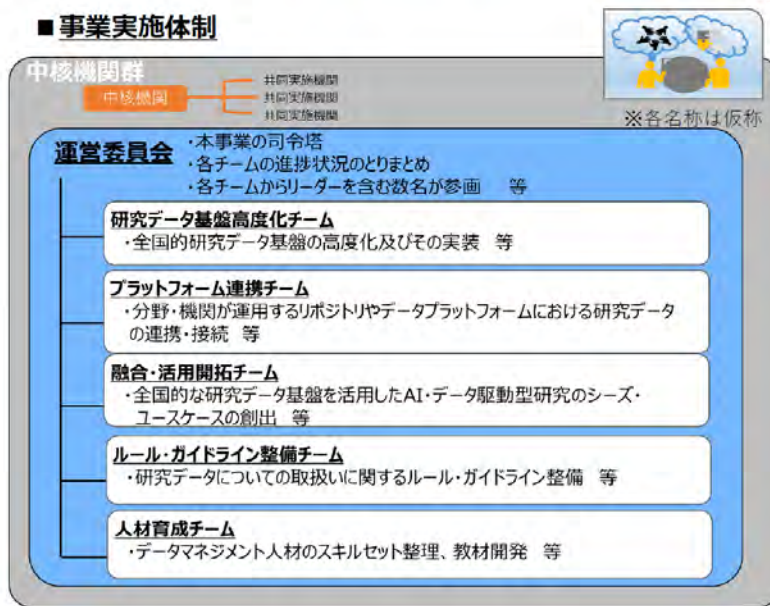
## 研究データエコシステムの構築に向けた具体的な取組イメージ（現行）

### ■全国的な研究データ基盤

全国的な研究データ基盤を中心に、オープン・アンド・クローズ戦略に基づいた研究データの管理・利活用が持続的に行われる仕組み（研究データエコシステム）の構築を目指す。そのために、全国的な研究データ基盤の構築・高度化・実装等や研究データ基盤の活用に係る環境整備を実施



### ■事業実施体制



### 中核機関群

研究データ基盤の構築・高度化・実装の中心的役割を担う中核機関が、複数の共同実施機関と連携した体制。中核機関群には、本事業の司令塔となる運営委員会及びそれぞれ推進すべき事業の内容に基づき中核機関群による機関から構成したチームを構成して事業推進

### ＜各チームのミッション＞

- ① **研究データ基盤高度化チーム**
  - ・全国的な研究データ基盤の高度化及びその実装
  - 【最低限求められる機能】
  - データマネジメントプラン（DMP）の内容に基づく研究データの取扱い（管理対象データの取捨選択やメタデータの付与等）を支援する機能
  - 研究データの出所や修正履歴等を管理し、対外的にその真正性を説明可能とすることを支援する機能
  - 研究者が管理する研究データを共有・公開した際、その研究データの利用状況や公開後の再利用の状況をモニタリングし、研究者にフィードバックする機能
  - 研究者が管理する研究データ・プログラム・解析環境等の情報を、再現性の検証や再利用可能な形でパッケージ化し公開することを支援する機能
  - 秘匿すべき情報が含まれる研究データについて、その管理をシステム上で適切に行うとともに、これらの研究データを秘匿したまま安全に解析可能とするための環境構築を支援する機能
- ② **プラットフォーム連携チーム**
  - ・分野・機関が運用するレジストリやデータプラットフォームにおける研究データの連携・接続
  - ・全国的な研究データ基盤上で統合的なメタデータ検索の実現
  - ・全国的な研究データ基盤との接続を希望する産学官の機関等に提示する要件等の整備
- ③ **融合・活用開拓チーム**
  - ・全国的な研究データ基盤を活用し、異なる分野間でのデータ連携を前提としたAI・データ駆動型研究のシーズ・ユースケースの創出
  - ・全国的な研究データ基盤のユーザー拡大に向けた周知・広報活動
- ④ **ルール・ガイドライン整備チーム**
  - ・研究データの効果的・効率的な活用のための機械可読データの統一化や標準化等や個人情報等の研究データを取り扱う上で留意すべきデータを含む研究データについての取扱いに関するルール・ガイドラインの整備
- ⑤ **人材育成チーム**
  - ・研究データ管理に必要なデータマネジメント人材も必要なスキルセット等の要件の整理
  - ・データマネジメント人材の育成に向けた教材開発やコンテンツ整備等