

研究データ基盤システムの利用実績(2)

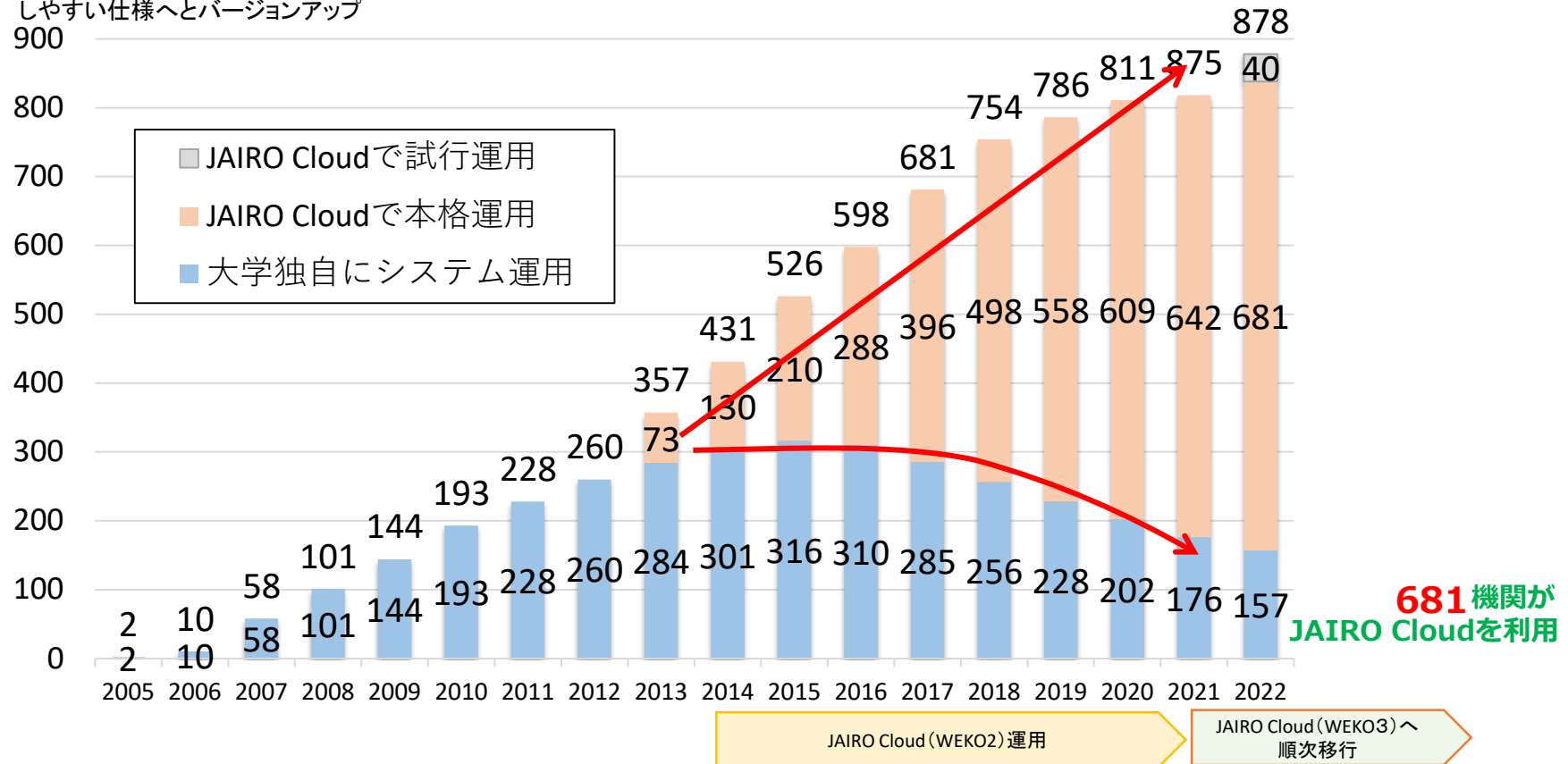
JAIRO Cloud

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へ順次移行途中で、

※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

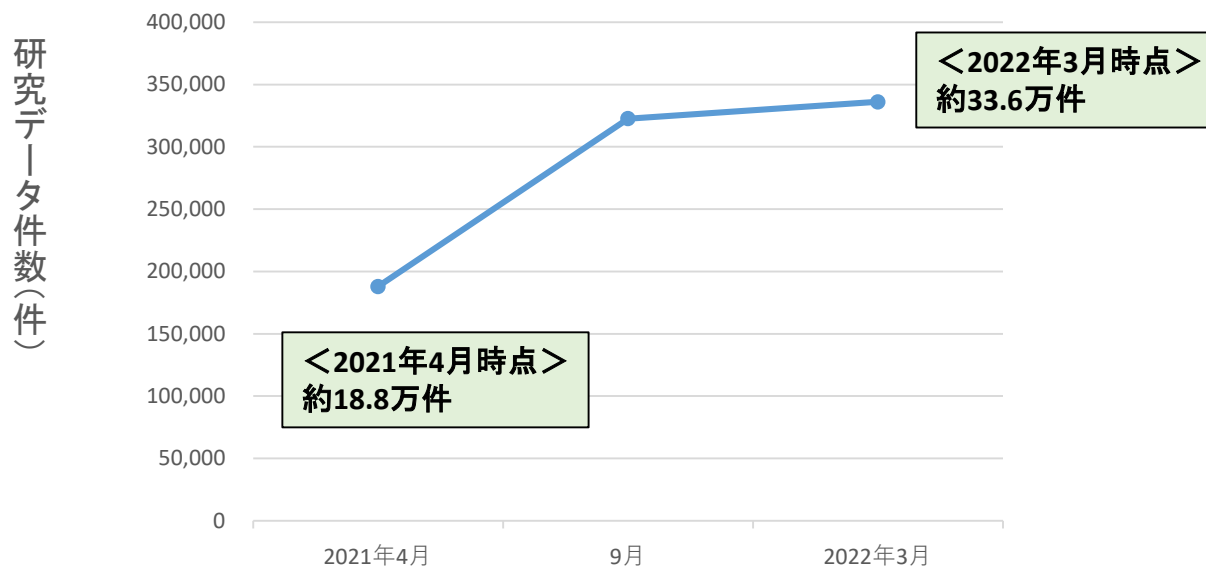


研究データ基盤システムの利用実績(3)

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。

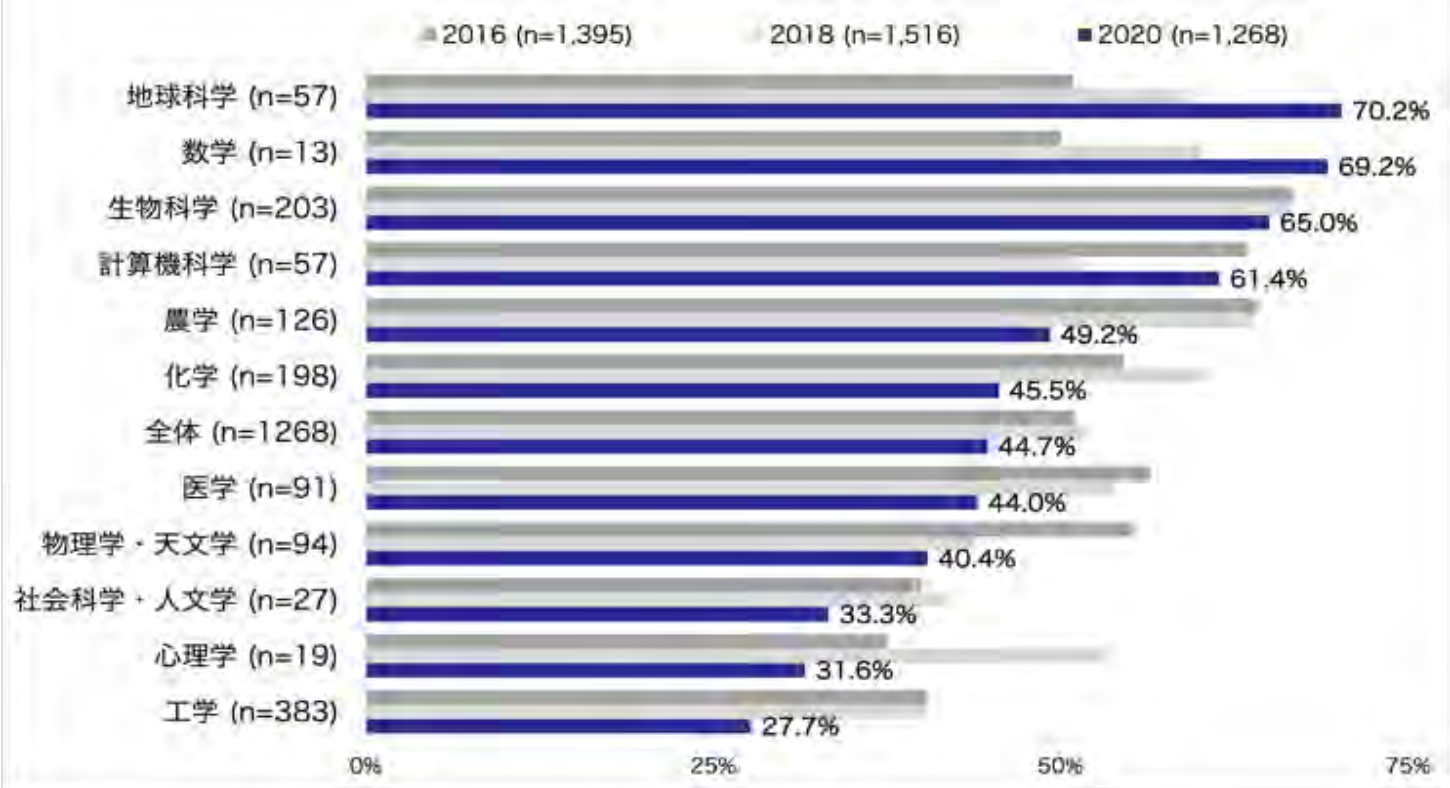
＜CiNii Research で検索可能な研究データ収載数＞



※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。

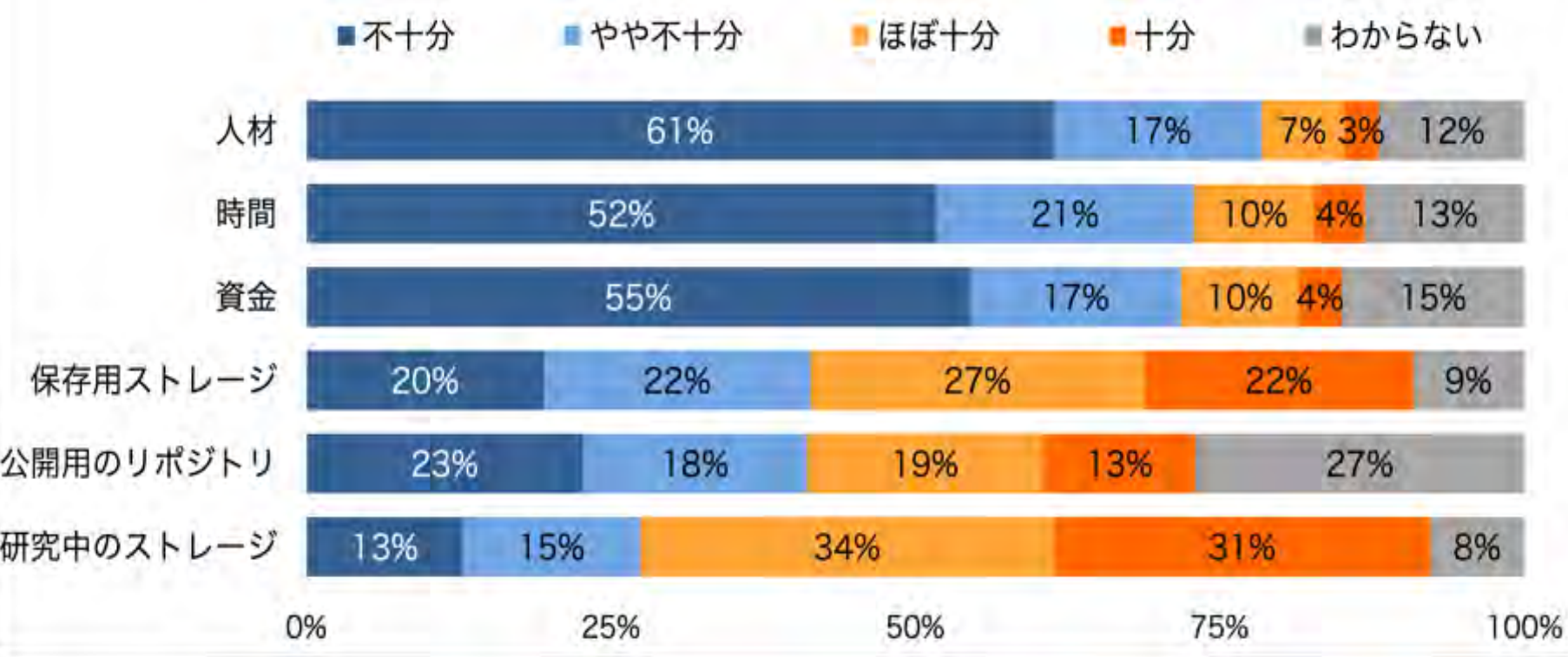
分野別データ公開経験



- 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

データの整備・公開資源の充足度

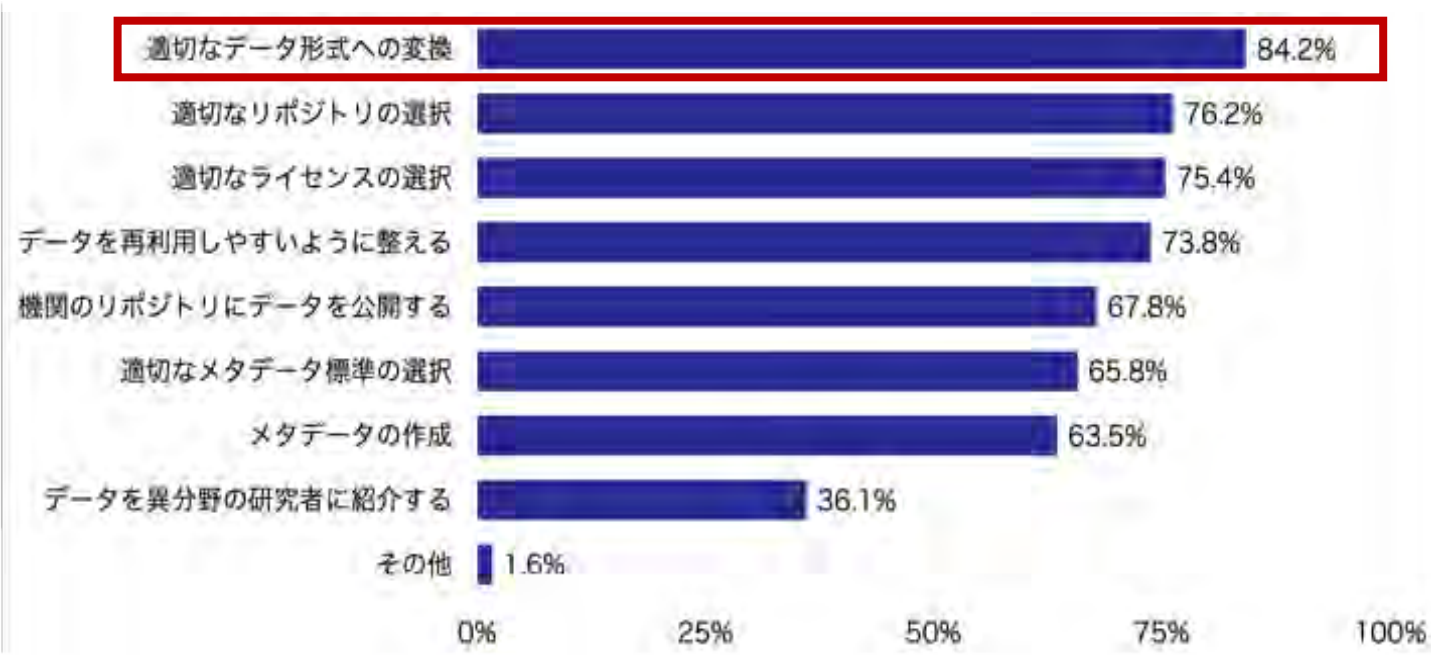
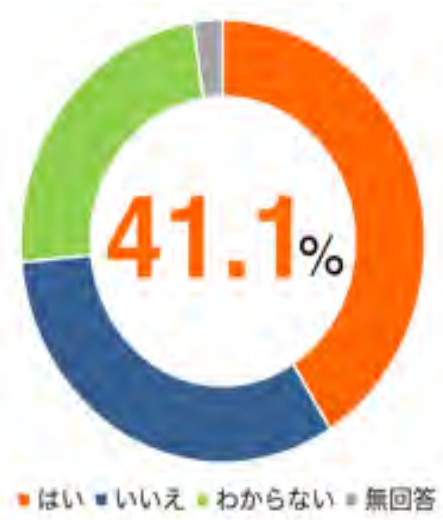
(n=1,188)



- 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いえない
- 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>)

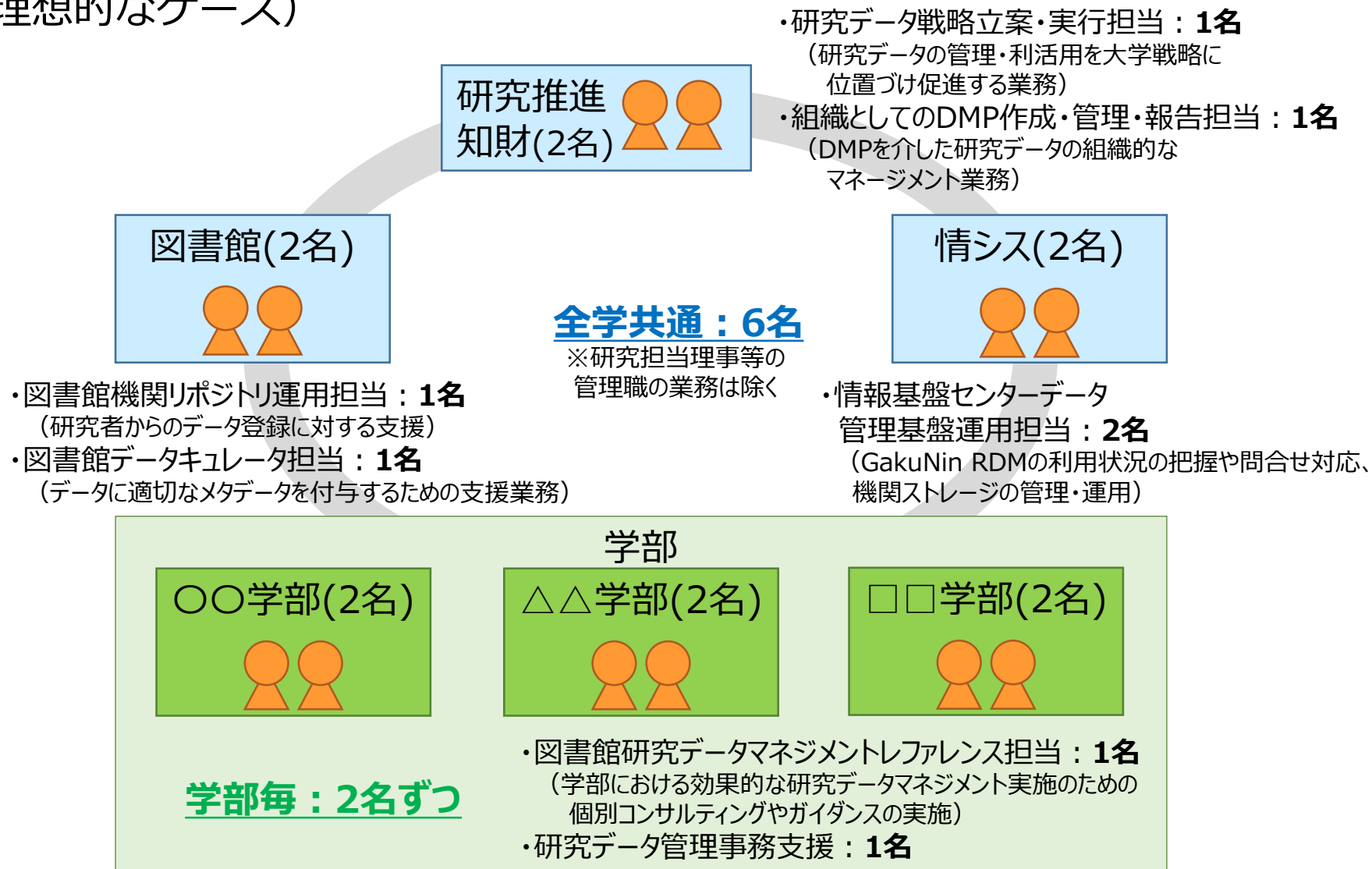
データの整備・公開・保存の依頼意思



- 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

研究データマネジメント業務の規模感（イメージ）

5,000人程度の大学の場合 （理想的なケース）



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ぎ**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
- ・各機関リポジ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

<事業スキーム>



中核機関群
(1件)

<実施体制>



我が国の研究力の飛躍的發展へ

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

研究データ基盤高度化チーム

NII Research Data Cloudを
7つの側面から機能拡張

NII

リーダ機関

研究データ基盤の機能実装

活用

コード付帯機能

データ・プログラム・解析環境のパッケージ化と流通機能を提供し、研究成果の再現性を飛躍的に向上

信頼

データプロビナンス機能

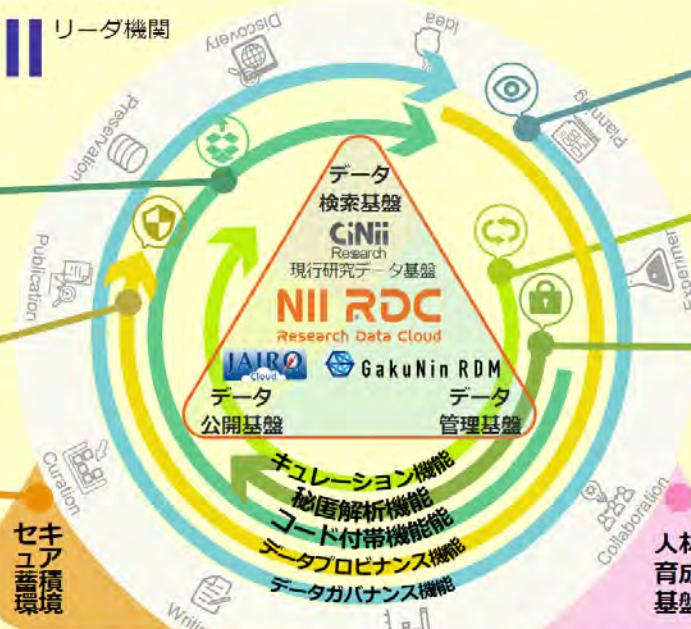
データの来歴情報の管理から利用状況を把握でき、データ公開へのインセンティブモデルを提供

蓄積

セキュア蓄積環境

安全で強固なデータの保存・保護機能を有する超鉄壁ストレージを提供し、機微な情報も安心して保全

セキュア蓄積環境



データガバナンス機能 管理

計画に基づきデータ管理等を機械的に支援し、DMPをプロジェクト管理に不可欠な仕組みへと変革

キュレーション機能 流通

専門的なキュレーションを実践できるエコシステムを構築し、データ再利用の促進に寄与

秘匿解析機能 保護

秘密計算技術で機微な情報も安心して解析できる環境の提供で、新しいデータ駆動型研究の世界を開拓

人材育成基盤 育成

RDMに必要なスキルを学ぶ環境を提供し、全ての研究者を新しい科学の実践者へと育成

プラットフォーム連携チーム

理化学研究所
RIKEN

リーダ機関

- ・ 機関内サービス等とNII RDCの連携機能の整理と設計
- ・ 計測機器等からの大量データを効果的に管理するための要件整理と機能開発
- ・ 管理対象となるメタデータの設計と実証
- ・ 関連する高度化機能との仕様調整と共同開発

融合・活用開拓チーム

東京大学
THE UNIVERSITY OF TOKYO

リーダ機関

- ・ 異なる分野間でのデータ活用やデータ連携に発展する取り組みを精査
- ・ 異なる分野間でのデータ活用やデータ連携に関する具体的なユースケースを創出
- ・ ユースケースをまとめたツールキットの作成とそれを用いた広報活動

ルール・ガイドライン整備チーム

名古屋大学
NAGOYA UNIVERSITY

リーダ機関

- ・ 研究データの活用に適した機械可読データの統一的な記述ルール設計
- ・ 研究データの公開に必要な要項や作業フローの整備
- ・ 研究データを適切に取扱うための指針のまとめ
- ・ 学内整備のための事例形成

人材育成チーム

大阪大学
OSAKA UNIVERSITY

リーダ機関

- ・ 人材育成を主とした研究データ管理体制の構築を推進する学内組織構築の事例形成
- ・ 研究データ管理人材に求められる標準スキルに関する検討
- ・ 研究データ管理人材育成のためのカリキュラムの作成、オンライン学習コースの整備

基盤の活用に係る環境整備

中核機関群の代表からなる運営委員会が全体を統括し研究データエコシステムの全国展開に向けて共同実施機関を随時拡大