https://doi.org/10.15108/stih.00287

ほらいずん

日本の研究機関における研究データ管理（RDM）の実践状況
－オープンサイエンスの実現に向けた課題と展望－

データ解析政策研究室 客員研究官 池内 有為、室長 林 和弘

【概　要】

2020 年にオープンアクセスリポジトリ推進協会（JPCOAR）と大学 ICT 推進協議会（AXIES）は、国内の大学や研究機関を対象とした研究データ管理（RDM）の取組状況に関するオンライン調査を実施した。科学技術・学術政策研究所（NISTEP）データ解析政策研究室は、結果データの提供を受けて二次分析を実施した。

352 件の回答を分析した結果、データポリシーを策定・検討している機関は 23.6%、RDM 体制を構築・検討している機関は 23.0% であった。RDM 体制のステークホルダーとして認識されていたのは、研究推進・協力系部門（61.9%）、図書館（43.2%）、情報系センター（31.5%）の順であった。RDM サービスのための情報インフラ整備を検討・対応している機関は 17.8%、データを長期保存するためのストレージを検討・提供している機関は 21.9% であった（この 2 問のみ n=297）。いずれの実施・検討率も、大学共同利用機関や研究開発法人の方が大学よりも高かった。大学の実施・検討率は、国立大学、私立大学、公立大学の順に高く、また、学部数が多い大学の方が高かった。RDM 体制の構築状況や RDM サービス提供経験の有無によって、データ公開の障壁や必要な能力の認識に差がみられた。

キーワード：研究データ管理（RDM），データポリシー，オープンサイエンス，研究データ公開，情報インフラ

1. はじめに

2021 年 6 月に閣議決定された「統合イノベーション戦略 2021」[1] において、「科学技術・イノベーション政策において目指す主要な数値目標」（主要指標）として、(1) 機関リポジトリ（Institutional Repository, IR）を有する全ての大学・大学共同利用機関法人・国立研究開発法人において、2025 年までに、データポリシーの策定率が 100%になること、(2) 公募型の研究資金の新規公募分において、2023 年度までにデータマネジメントプラン（DMP）[注1] 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が 100% になることが掲げられた（p.81）。すなわち、日本の大学や研究機関には、データポリシーの策定、及び研究データを適切に保存し、必要に応じて第三者に開示・公開するための研究データ管理（Research Data Management, RDM）体制の構築が求められている。

こうした状況をふまえて、オープンアクセスリポジトリ推進協会（JPCOAR）と大学 ICT 推進協議会（AXIES）は、2020 年に国内の大学や研究機関における RDM の取組状況を把握するためのオンライン質問紙調査を実施した[2]。科学技術・学術政策研究所（NISTEP）データ解析政策研究室は、今後の学術情報政策立案に資することを目的として、調査結果から回答者の情報を除いたデータの提供を受けて分析を行った。本稿では、まず調査の概要を示した後に主要な分析結果を報告する。

2. 調査の概要

調査は 2020 年 11 月から 12 月にかけて、JPCOAR 研究データ作業部会と AXIES-RDM 部会が共同で

注 1　研究のために収集・作成するデータをどのように管理するか、取り扱いや整備・保存・公開に関する計画を記した書類を指す。

（2022.3.22 公開）

実施した。国内の大学や研究機関を対象として、JPCOAR に所属する 657 機関、AXIES に所属する 133 機関にメールで参加を呼びかけた（それぞれの所属機関は一部重複している）。

有効回答 352 件のうち、大学は 325 件（92.3%）、研究機関等は 16 件（4.6%）、その他は 11 件（3.1%）であった。大学は国立大学が 67 件（19.0%）、公立大学が 42 件（11.9%）、私立大学が 216 件（61.4%）であり、研究機関は研究開発法人と大学共同利用機関がそれぞれ 8 件（2.3%）ずつであった。その他には、高等専門学校や独立行政法人等が含まれる。

質問数は全 46 問であり、①メールアドレス、②基礎情報（回答部署等）、③機関構成員のニーズの把握状況、④ RDM 体制の構築状況、⑤ RDM サービスの実施状況、⑥情報インフラの整備状況、⑦ JPCOAR についての 7 セクションである。うち、⑥は AXIES 参加機関と JPCOAR 参加機関でフォームを分けて回答を収集したため、有効回答は 297 件（全有効回答の 84.4%）であった。

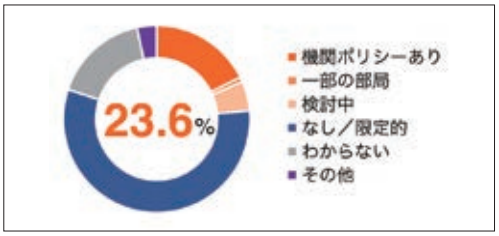
3. 分析結果

ここでは、(1) データポリシーの策定状況、(2)RDM 体制の構築状況、(3) 情報インフラの提供状況、(4)IR によるデータの公開状況と障壁、(5)RDM サービスの展開と必要な能力に関する結果を示す。なお、以下では回答者数を「n」で表す。

3.1　データポリシーの策定状況

「貴機関では、研究データ管理に関する何らかのデータポリシーは定められていますか」という質問に対する回答の集計結果を図表 1 に示す。質問文中では「データポリシー」を定義していないが、『統合イノベーション戦略 2021』[1] で示されたような RDM 全般に関するデータポリシーを想定しており、データの 10 年保存[3] に関する規程のみを定めている場合と区別している。具体的には、選択肢として「機関としてのポリシーを制定した（以下、「機関ポリシーあり」）」と「ポリシーは定められていない／限定的なポリシー（例：データ保存ポリシー）が定められている（以下、「なし／限定的」）」を提示している。

図表 1　データポリシーの策定状況（n=352）



図表 1 では、「機関ポリシーあり」（17.6%）、「一部の部局で定められている（以下、「一部の部局」）」（0.9%）、「機関全体でのポリシーを検討中である（以下、「検討中」）」（5.1%）の合計（23.6%）を数値として示した。「なし／限定的」（56.5%）の比率が最も高く、「わからない」（16.5%）という回答もみられた。

図表 2 に、機関の種類ごとにデータポリシーの策定状況を集計した結果を示す。配列は、「機関ポリシーあり」、「一部の部局」、「検討中」の合計が多い順とした。

ポリシーの策定・検討率は、回答者数は少ないものの研究開発法人（75.0%）と大学共同利用機関（62.5%）が高く、次いで国立大学（38.8%）、私立大学（19.0%）、その他（18.2%）、公立大学（7.1%）の順であった。

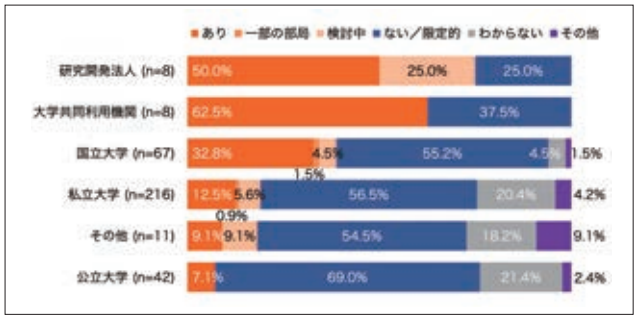
図表 3 に、大学のみを対象として、大学の規模（学部数）ごとにデータポリシーの策定状況を集計した結果を示す（n=325）。配列は、規模の昇順とした。

ポリシーの策定・検討率は、おおむね規模が大きいほど高く、8 学部以上の大学（40.0%）、5〜7 学部（27.9%）、2〜4 学部（15.7%）、単科大学（16.3%）であった。逆に、規模が小さい大学ほどポリシーが「ない／限定的」の選択率が高かった。

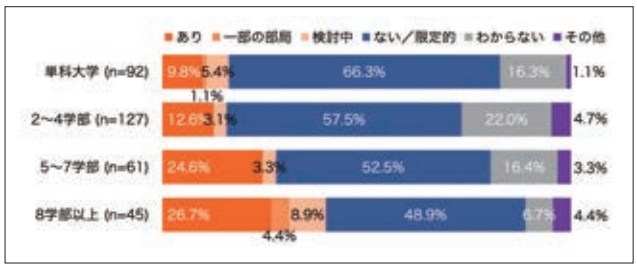
3.2　研究データ管理（RDM）体制の構築状況

「貴機関では、機関全体での研究データ管理体制構築に向けた何らかの取り組みが始まっていますか」という質問について、集計結果を図表 4 に示す。図表 4 には、何らかの取組を始めている機関の比率、すなわち「体制が構築され、機関として実施してい

図表 2　機関別：データポリシーの策定状況（n=352）



図表 3　大学の規模別：データポリシーの策定状況（n=325）



る」（8.5%）、「一部の部局で議論・実施されている」（8.8%）と「検討のための委員会、ワーキンググループ等が組織されている」（5.7%）の合計（23.0%）を数値で示した。「必要性は認識しているが、具体的な動きはない」（58.5%）の比率が最も高く、「わからない」（18.5%）という回答もみられた。

図表 5 に、機関の種類ごとに RDM 体制の構築状況を集計した結果を示す。配列は、「機関として実施」、「一部の部局で議論・実施」、「ワーキンググループ（WG）等を組織化」の合計が多い順とした。

機関として実施している比率が高かったのは研究開発法人と大学共同利用機関（いずれも 25.0%）であり、次いで国立大学（16.4%）、その他（18.2%）、私立大学（5.1%）、公立大学（4.8%）の順であった。

図表 6 に、大学のみを対象として、大学の規模ごとに RDM 体制の構築状況を集計した結果を示す（n=325）。配列は、規模の昇順とした。

「機関として実施」、「一部の部局で議論・実施」、「WG 等を組織化」の比率の合計は、規模が大きいほど高かった。ただし、「機関として実施」の比率は 8 学部以上（2.2%）が低い点が機関レベルのデータポリシーの策定率とは異なっていた。

また、「研究データ管理体制に関する議論で、ステークホルダーになり得る部署、または既に関与している部署はどこでしょうか」という複数選択方式の質問について、図表 4 に示した機関レベルの RDM 体制の有無とクロス集計を行った。結果を図表 7 に示す。

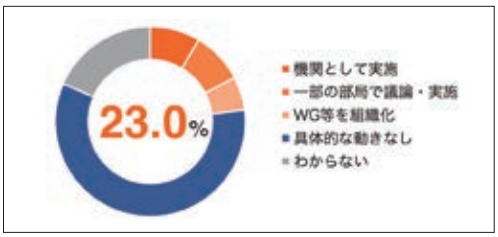
ステークホルダーとして選択率が高かったのは、「研究推進・協力系部門」（合計 61.9%）、「図書館」（同 43.2%）、「情報系センター」（同 31.5%）の順であった。RDM 体制ありの機関の回答は、順位は変わらないものの「研究推進・協力系部門」（80.0%）の選択率が有意に高く（$p<0.05$）、「図書館」、「情報系センター」、「知財系部門」は選択率が低かった。

3.3　情報インフラの提供状況

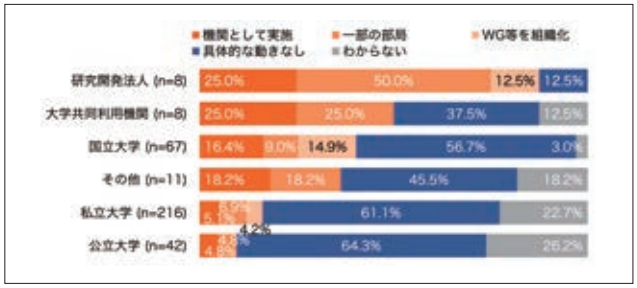
ここでは AXIES の所属機関を対象とした、研究データ管理に関連する情報インフラ、すなわちストレージや公開用リポジトリ、データ分析基盤等の提供状況に関する質問の結果を示す。まず、「現在、機関における研究データ管理サービスの情報インフラ整備はどのように検討が進められていますか」という質問の集計結果を図表 8 に示す。図表 8 には、何らかの検討や対応を進めている比率、すなわち「機関レベルで検討・対応が進められている」（10.4%）と「部局ごとに検討・対応が進められている」（7.4%）の合計（17.8%）を数値で示した。最も選択率が高かったのは「特に検討・対応は行われていない」（43.1%）、次いで「研究者個人に委ねられている」（20.9%）であり、「わからない」（17.8%）や「詳細は不明」（0.3%）という回答もみられた。

図表 9 に、機関の種類ごとに情報インフラ整備の検討状況を集計した結果を示す。配列は、「機関レベ

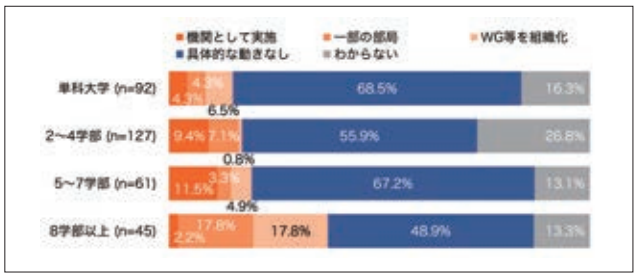
図表 4　RDM 体制の構築状況（n=352）



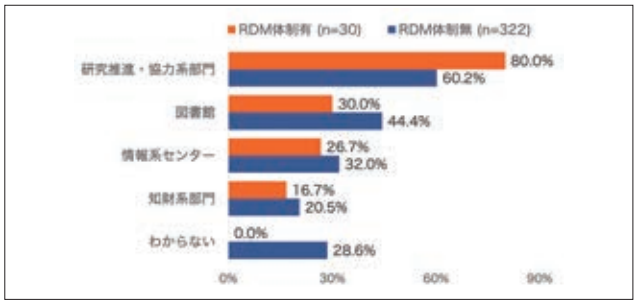
図表 5　機関別：RDM 体制の構築状況（n=352）



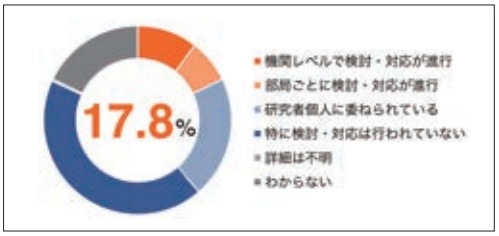
図表 6　大学の規模別：RDM 体制の構築状況（n=325）



図表 7　RDM 体制構築の有無とステークホルダー（n=352）



図表 8　RDM サービスに関する情報インフラ整備の検討状況（n=297）



ル」と「部局ごと」の合計が多い順とした。

データポリシーや RDM 体制と同様に、研究開発法人や大学共同利用機関は相対的に情報インフラ整備の検討・対応が進められており、「なし」の比率が低かった。また、他の調査結果と同様に国立大学、私立大学、公立大学の順に検討・対応が進められていた。

図表 10 に、大学のみを対象として、大学の規模ごとに情報インフラ整備の検討状況を集計した結果を示す（n=273）。配列は、規模の昇順とした。

規模が大きい大学ほど情報インフラ整備の検討・対応が進められており、「なし」の比率が低かった。また、研究者個人に委ねられているという回答は、規模が小さい単科大学（23.5%）と 2～4 学部の大学（23.7%）で比較的高く、規模が大きい 5～7 学部（17.4%）や 8 学部以上の大学（17.9%）と差がみられた。

続いて「研究データを長期的（5 年以上）に保存するためのストレージ（オンプレミス[注 2] 又は商用クラウド）を用意していますか」という質問について、集計結果を図表 11 に示す。図表 11 には、「オンプレミスで提供している」（8.1%）、「商用クラウドで提供している」（6.4%）、「上記のいずれか、あるいは両方で提供を検討している」（7.4%）の合計（21.9%）を数値で示した。ストレージについては、「提供する予定はない」（47.1%）、及び「わからない」（31.0%）の選択率が高かった。

図表 12 に、機関の種類ごとにデータの長期保存用ストレージの提供状況を集計した結果を示す。配列は、「オンプレミス」、「商用クラウド」、「提供を検討中」の合計が多い順とした。

大学の比率はここでも相対的に低く、他の調査結果と同様に国立大学、私立大学、公立大学の順であった。

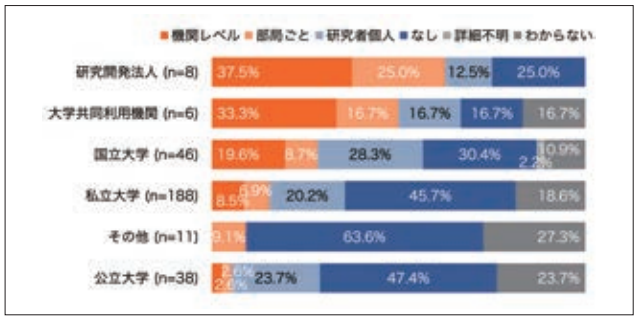
図表 13 に、大学のみを対象として、大学の規模ごとにデータの長期保存用ストレージの提供状況を集計した結果を示す（n=273）。配列は、規模の昇順とした。

規模が大きい大学ほど提供・検討している比率が高かった（ただし、単科大学の方が 2～4 学部の大学よりもやや高かった）。逆に規模が小さい大学ほど「提供予定なし」の比率が高かった。

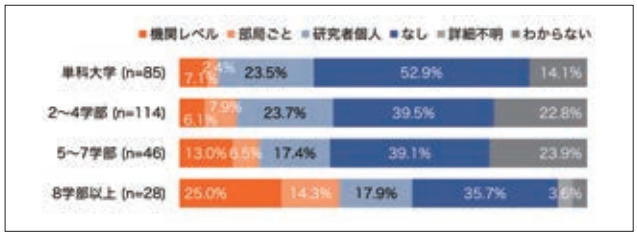
3.4　機関リポジトリ（IR）によるデータの公開状況と障壁

「貴機関の機関リポジトリで研究データを公開している事例はありますか」という質問に対して、「ある」と回答したのは 88 件（25.0%）、「ない」は 252 件（71.6%）、「わからない」は 12 件（3.4%）であった。機関の種類ごとに集計すると、「ある」の選択率が高

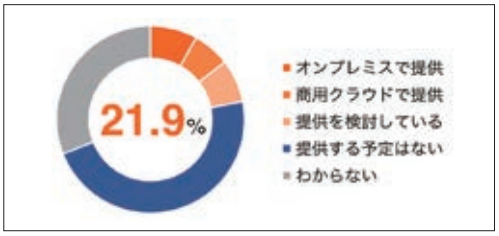
図表 9　機関別：RDM サービスに関する情報インフラ整備の検討状況（n=297）



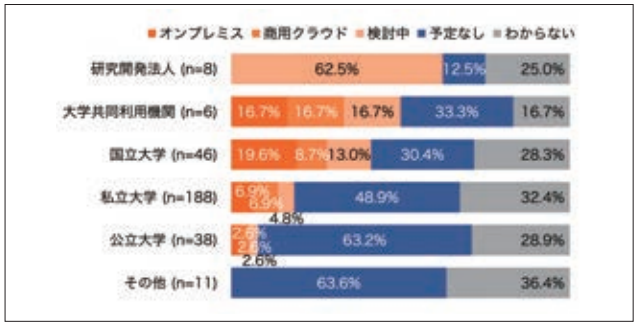
図表 10　大学の規模別：RDM サービスに関する情報インフラ整備の検討状況（n=273）



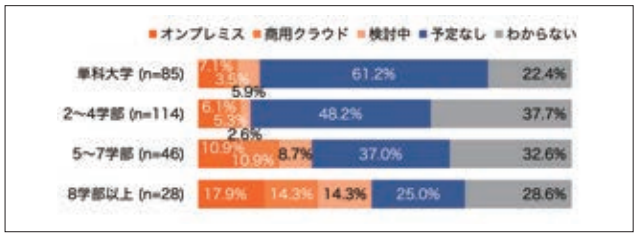
図表 11　研究データを長期保存するためのストレージの提供状況（n=297）



図表 12　機関別：研究データを長期保存するためのストレージの提供状況（n=297）



図表 13　大学の規模別：研究データを長期保存するためのストレージの提供状況（n=273）



注 2　サーバやソフトウエアなどの情報システムを機関内に設置して運用する方式。

い順に研究開発法人（50.0%）、国立大学（34.3%）、大学共同利用機関（25.0%）、私立大学（24.1%）、その他（18.2%）、公立大学（11.9%）の順であった。また、大学の規模別に集計すると（n=325）、8 学部以上（42.2%）、5〜7 学部（24.6%）、2〜4 学部（22.0%）、単科大学（19.6%）であった。

続いて「機関リポジトリでの研究データ公開にあたり、課題や障壁となり得ることをお聞かせください」という複数選択方式の質問について、IR による公開事例が「ある」「ない」「わからない」機関ごとにクロス集計した結果を図表 14 に示す。

最も選択率が高かったのは「公開に当たってのマンパワーが足りない」（64.8%）、次いで「適切なライセンス・利用条件がわからない」（53.4%）、「機関リポジトリの運用規程等が未整備である」（47.2%）の順であった。IR による公開事例の有無によって選択率に有意な差がみられたのは、「研究データのフォーマット形式がわからない、扱えない」、「研究データの利用形態がわからない」、「機関リポジトリの運用規程等が未整備である」、「適切なライセンス・利用条件がわからない」の 4 項目であり、いずれも「ない」機関の方が選択率が高かった。

3.5　RDM サービスの展開と必要な能力

「研究データ管理サービスを展開するにあたり、自機関の研究者から要望が高いと思われるサービスをお聞かせください」という複数選択方式の質問について、機関レベルの RDM 体制の有無でクロス集計した結果を図表 15 に示す。

最も選択率が高かったのは「わからない」（全体の 56.0%）であり、RDM 体制を構築済みの機関であっても 15 機関（50%）が選択していた。サービスについては「研究データのストレージ提供」（23.3%）、次いで「データ公開プラットフォーム（IR 等）の提供」（全体の 22.4%）、「研究データの知財管理支援」（全体の 20.2%）の順であった。「データキュレーション[注3]支援」（全体の 14.8%）は最も選択率が低かったが、RDM 体制を有する機関（n=30）に限ると、「データキュレーション支援」（30.0%）はストレージ提供に次ぐ 2 位であった。

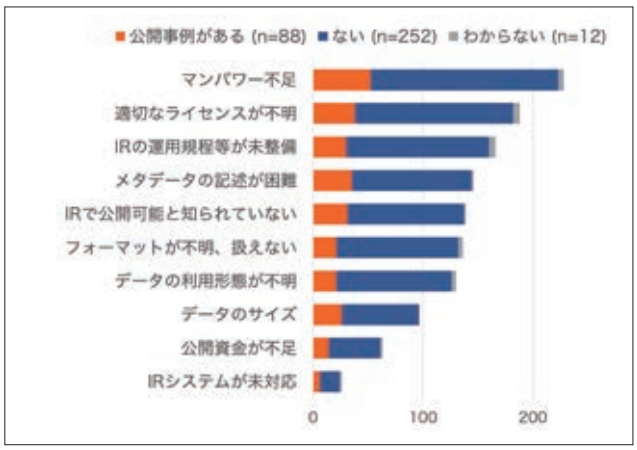
また、RDM に関する支援を実施する際に必要と思われる知識やスキルを尋ねた質問でも、「データキュレーションのスキル」（49.4%）は「ICT スキル」（52.8%）や「知財等に関する知識」（49.4%）に次ぐ 3 位であったが、RDM 体制を構築している機関（n=30）に限ると 1 位（60.0%）であった。つまり RDM 体制を有する機関では、データキュレーションに対するニーズが高いと考えられており、対応するための知識やスキルが必要であると認識されていた。

4. RDM の現状と課題

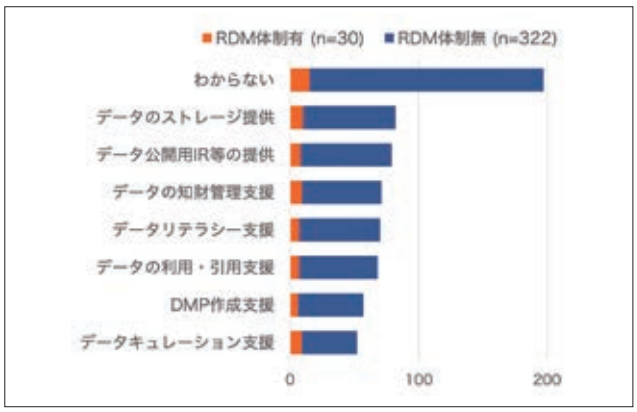
本稿のまとめとして、(1) データポリシーと RDM 体制の構築状況、(2) 物的資源と人的資源の状況、(3) RDM 体制を構築している機関から得られた示唆について述べる。

現状では、データポリシーの策定や RDM サービスは、一部の大学や研究機関でのみ実施されていた。おおむね研究開発法人や大学共同利用機関の実施率が高く、次いで国立大学、私立大学、公立大学の順であった。大学の規模別にみると、大学の規模が大きいほど実施率が高い傾向がみられた。ただし、RDM 体制を機関全体で実施している比率については「8 学部以上」の大学で低く、規模が大きい大学では全学的な体制を構築することが困難である可能性が示唆された。

図表 14　機関リポジトリによるデータ公開の課題や障壁（n=352）



図表 15　研究者からの要望が高いと思われるサービス（n=352）



注 3　研究データを管理・公開・共有・長期保存するために、データを整備したりデータに関する情報（メタデータ）を付与したりすること。

RDMのための情報インフラを提供している大学や研究機関は2割程度であった。RDMサービスのための情報インフラは43.1%が検討や対応をしておらず、研究データの長期保存用ストレージは47.1%が提供する予定はないとしていた。国立情報学研究所（NII）によるNII研究データ基盤[4]は、「統合イノベーション戦略2021」[1]において研究データの管理・利活用のための中核的なプラットフォームとして位置づけられている。こうした効率的かつ安心安全な情報インフラが広く普及し、オールジャパンのRDM体制が整備されることは、研究者や管理者のみならず、その成果がもたらされる社会全体にとってもメリットが大きいと考えられる。

RDM体制構築のためのステークホルダーとしては、研究推進・協力系部門、図書館、情報系センターの選択率が高かった。既にRDM体制を構築している、あるいは検討中の機関に限ると、研究推進・協力系部門の選択率は81.5%にのぼった。今後、研究推進・協力系部門のリサーチ・アドミニストレーター（URA）がRDM支援を行っていくことが考えられるが、URAは有期雇用である場合がほとんどであることから、長期的な視点でRDMサービスを検討し、経験を蓄積するために長期にわたる人材配置が必要であろう。また、組織としてRDMを複数の部署で円滑に進めていくためには、全体を統括する意思決定者、及び橋渡し人材が必要であると考えられる。

最後に、データポリシーの策定やRDM体制の構築を効率的に進めていくためにはどうすればよいのか、これからRDMに取り組む機関と既に実践している機関に分けて検討したい。IRによるデータ公開を行っていない機関では、研究データのフォーマットやライセンス等、基本的な事項が課題であると認識している回答者の比率が高かった。JPCOAR研究データ作業部会は、RDMの基本的な教材として、支援対象者向けに「研究データ管理サービスの設計と実践」（第2版）[5]を、研究者向けに「研究者のための研究データマネジメント」[6]を無料で公開している。また、AXIES研究データマネジメント部会は「大学における研究データポリシー策定のためのガイドライン」[7]を策定し、関連イベントを開催している。これからRDMに取り組む機関にとっては、こうした教材やイベントの活用が有効であろう。一方、既にRDMを実施している機関では、RDMサービスの中でもデータキュレーション支援のニーズが高く、対応するための能力が支援者に必要であると認識されていた。先行事例の知見を共有するとともに、データキュレーションの多岐にわたる内容を整理し、必要な知識や技術は何か、誰がどのような支援を担うのか、人材育成の仕組みとキャリアパスをどう設計するのか、現状をふまえた議論と対応が必要であろう。

謝辞

本二次分析に当たり、オープンアクセスリポジトリ推進協会（JPCOAR）及び大学ICT推進協議会（AXIES）から「国内機関における研究データの取り組み状況調査」データの提供を受けた。また、JPCOARとAXIESの皆様には情報の提供や議論に御協力いただいた。心より御礼申し上げる。

参考文献・資料

1) 内閣府 . 統合イノベーション戦略2021：本文 . 2021, 113p.
https://www8.cao.go.jp/cstp/tougosenryaku/togo2021_honbun.pdf, (accessed 2021-12-14).
2) 南山泰之 , 結城憲司 , 田邉浩介 , 安原通代 . 2020年度RDM事例形成プロジェクト中間報告書
https://jpcoar.repo.nii.ac.jp/records/2000003
3) 日本学術会議 . 科学研究における健全性の向上について（回答）. 2015, 35p.
https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-k150306.pdf, (accessed 2021-12-14).
4) “NII研究データ基盤（NII Research Data Cloud）の概要”. オープンサイエンス基盤研究センター .
https://rcos.nii.ac.jp/service/ (accessed 2021-12-14).
5) オープンアクセスリポジトリ推進協会（JPCOAR）研究データ作業部会 . 研究データ管理サービスの設計と実践：第2版 . 2021, [PowerPoint]. https://jpcoar.repo.nii.ac.jp/records/607, (accessed 2021-12-14).
6) オープンアクセスリポジトリ推進協会（JPCOAR）研究データ作業部会 . 研究者のための研究データマネジメント . 2020, [PowerPoint]. https://jpcoar.repo.nii.ac.jp/records/294, (accessed 2021-12-14).
7) 大学ICT推進協議会（AXIES）. 大学における研究データポリシー策定のためのガイドライン . 2021, 70p.
https://rdm.axies.jp/sig/70/, (accessed 2021-12-14).