

政府関係機関事業紹介

研究データ基盤の運用開始と未来に向けて

国立情報学研究所 オープンサイエンス基盤研究センター

1. COVID-19と研究データ基盤環境の整備

国立情報学研究所（NII）では、大学や研究機関に向けての新しいサービス「NII Research Data Cloud（NII RDC）」の開発と運用を進めています。本誌の2019年度No.1では、NII RDCの開発に至った背景を紹介しました。No.2からNo.4では、NII RDCを構成する管理基盤のGakuNin RDM、検索基盤のCiNii Research、公開基盤WEKO 3をそれぞれ紹介しました。まだご覧になっていない方は、本稿と合わせてお読み頂けますと幸いです。

NII RDCを紹介する最後の記事が掲載されたところから、日本でもCOVID-19が猛威を振るい始めました。大学教育だけではなく、研究の在り方にも大きな変革が必要とされた一年でした。研究の側面では、世界が一丸となってCOVID-19に関するデータを共有し、人類の共有財産とすることで、この危機的な状況から脱しようとする動きを、我々は見てきました。そうした取組みが欧米を中心にいち早く広がったのは、単に共有する成果が多かったからではありません。長年にわたって、政府機関や研究費助成機関がオープンアクセスやオープンサイエンスに積極的に取り組んできた歴史があるからだと推測されます。その思想や行動がパンデミックの中で世界的に波及し、研究成果の公開や共有の在り方について、多くの研究者の意識を変えるきっかけになっています。ポストコロナ時代の新しい研究者の常識は、これまでに増してオープンサイエンスを支持するものになることは間違いありません。今後の科学技術政策は、さらにそれを後押しすることでしょう。研究データは再利用可能な形で適切に管理されると同時に、できる限り他者と共有することにより効率性と透明性を高めていくことが、研究者の取り組むべき重要な責務となってくる時代を迎えようとしています。

研究データの再利用性を最大化するためには、FAIR（Findable, Accessible, Inter-operable and

Reusable）データ原則が求めるように、公開前も含めて研究データが再利用され易い適切な状態で管理され、共有される必要があります。2018年の欧州委員会の報告によると、データがFAIR原則に準拠していないために、その共有や再利用が促進されないことの損失は、学術分野だけでも年間102億ユーロに上ると試算されています。ストレージやライセンス費用、研究費の重複支出など5項目への影響が算出根拠となっています。学際的研究や産学連携などの機会的損失も加味すると、さらに年間169億ユーロの損失が上乗せされます。

この経済損失の大きさに増して注視すべきことは、研究中のデータが適切に管理されずにFAIR原則に準拠していない現状が、新しい学術的活動を創成することへの阻害要因となっていることです。産学連携や学際的な研究に発展する機会を失っていることは、大学経営の立場から見逃すべきではありません。これらの不要かつ継続的な経済損失を避けるためには、研究者が必要とするデータ管理のための環境提供に、機関としても適切に投資することが肝要です。それはまさに、我々がコロナ禍で学んだ、オープンサイエンスを軸とした新しい科学のあり方を支援することに他なりません。新しくかつ効率的な学術研究スキームが実現可能な環境を整えていくためには、研究者だけではなく、情報基盤センター、研究担当事務、大学経営者のすべてが協調して課題解決にあたることが重要とされています。

2. NII Research Data Cloudの運用

NII RDCを構成する3つの基盤は、2020年度内に本格的な運用を開始します。新しいサービスとなるGakuNin RDMも、本誌が発行される頃には本格運用に突入しています。これまで進めてきた実証実験には22機関が参加し、様々なフィードバックを受けてきました。要望に基づく段階的な機能拡張を進めつつ、2021年1月からは日中365日の

監視体制での運用を開始しました。年度が変わる2021年4月からは、24時間体制での運用に切り替えます。

利用機関のコミュニティの育成にも、これからはさらに力を入れていく予定です。まず必要なのは、大学や研究機関が、GakuNin RDMを機関として採用するために必要とする機能を気軽に提案できる環境作りです。その仕様の詳細や開発の優先順位を、皆さんと共に決めることができる仕組みも用意します。

GakuNin RDMを導入する際には、学内の体制やポリシーも同時に準備する必要があります。主に基盤センターの職員から構成される大学ICT推進協議会（AXIES）の研究データマネジメント部会では、必要となる学内活動の情報共有を行っています。主に図書館員から構成されるオープンアクセスリポジトリ推進協会（JPCOAR）の研究データ作業部会では、研究者や支援者が研究データ管理を学ぶためのトレーニングコースを提供しています。それぞれのコミュニティにも参加して頂きながら、皆さまと共に学内における研究データ管理の仕組みを醸成していくことが、我々の使命です。

3. NII Research Data Cloudのさらなる挑戦

冒頭のCOVID-19とFAIRデータ原則に関する例でもふれたように、我々には、オープンサイエンスや研究データ管理推進への流れを、如何に機関としての研究力強化に繋げていくかという発想の転換が求められています。NII RDCとしても、研究推進のために研究者と機関の両側面から必要とされる機能を強化していくために、次期システムの準備を進めています（参照 図1）。

その一環として挙げられるのが、データガバナンス機能の研究開発です。FAIRデータ原則を満たす第一歩として、研究データ管理計画（DMP）というものがあります。一般的にDMPは、研究費助成機関に要請されて提出する書類という認識が多

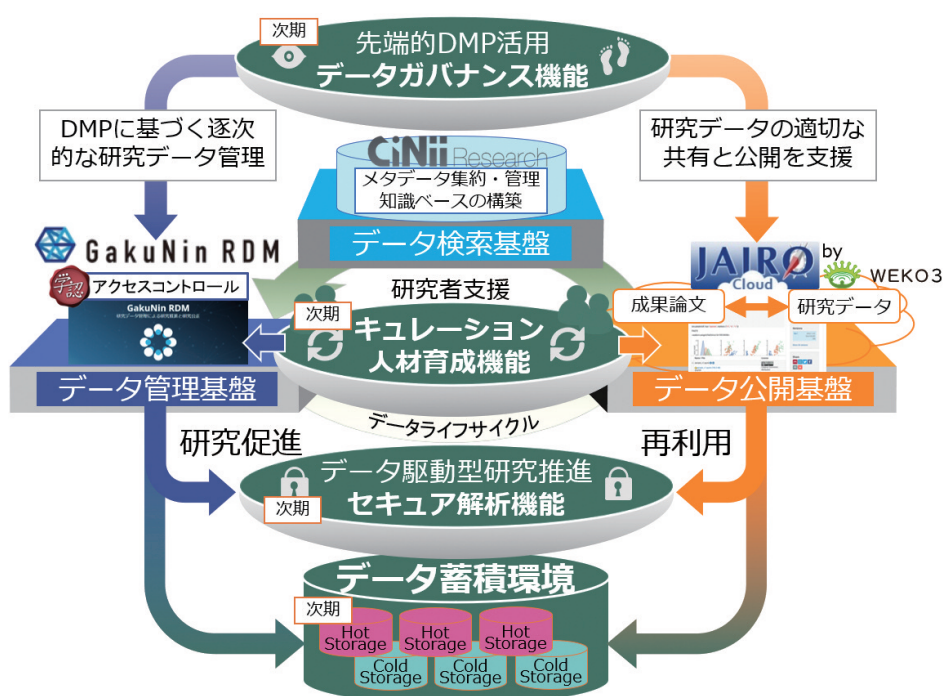


図1 次期NII Research Data Cloudの概略図

いと思いますが、海外ではこれを研究推進に積極的に活用しようとしています。その良い例が、豪州にあります。

豪州では、研究費助成機関から要請されるDMP自体は非常にシンプルです。その代わりに、政府から研究機関に対して、研究データ管理を推進する責務を課しています。先進的な大学では、研究データ管理とDMPを作成するツールをうまく連動させ、研究者が自然と研究データ管理を実践できる仕組みを提供しています。さらに、学内の研究推進室がDMPを活用し、共同研究やさらなる外部資金獲得を支援するという取り組みもあります。DMPを中心に、研究者がツールを使いこなし、学内部局との連携も促進されているという理想的な事例です。データガバナンス機能の開発チームは、こうした海外のグッドプラクティスをNII RDCにも取り込むべく、仕様の検討を進めています。

その他にも、データ駆動型研究を促進するためのセキュア解析機能、FAIRデータ原則の準拠に必要なデータキュレーションを支援する機能などの検討も同時に進めています。これらの新機能の開発を通して、学内の多面的な要望にかなうように、NII RDCのさらなる発展を目指します。今回は、紙面の都合上詳細な説明ができませんが、新しい機能の提供が具体化してきましたら、改めて本誌にて紹介させて頂きたいと思います。