

NII Today

91
Mar. 2021

National Institute of Informatics News

Feature

NII Research Data Cloud 本格始動へ

オープンサイエンスを支える研究データ基盤

NII Interview

次世代研究データ基盤
「NII RDC」への期待
オープンサイエンスを推進し、
イノベーションを促すために

橋爪 淳氏 [文部科学省研究振興局 参事官 (情報担当)]

座談会 NII×名古屋大学

NII研究データ基盤「NII RDC」が
いよいよ始まる!

山地一禎 [NIIオープンサイエンス基盤研究センター
センター長]

研究データ管理基盤

「GakuNin RDM」の本運用がスタート

日本の学術機関のデータ公開を支える

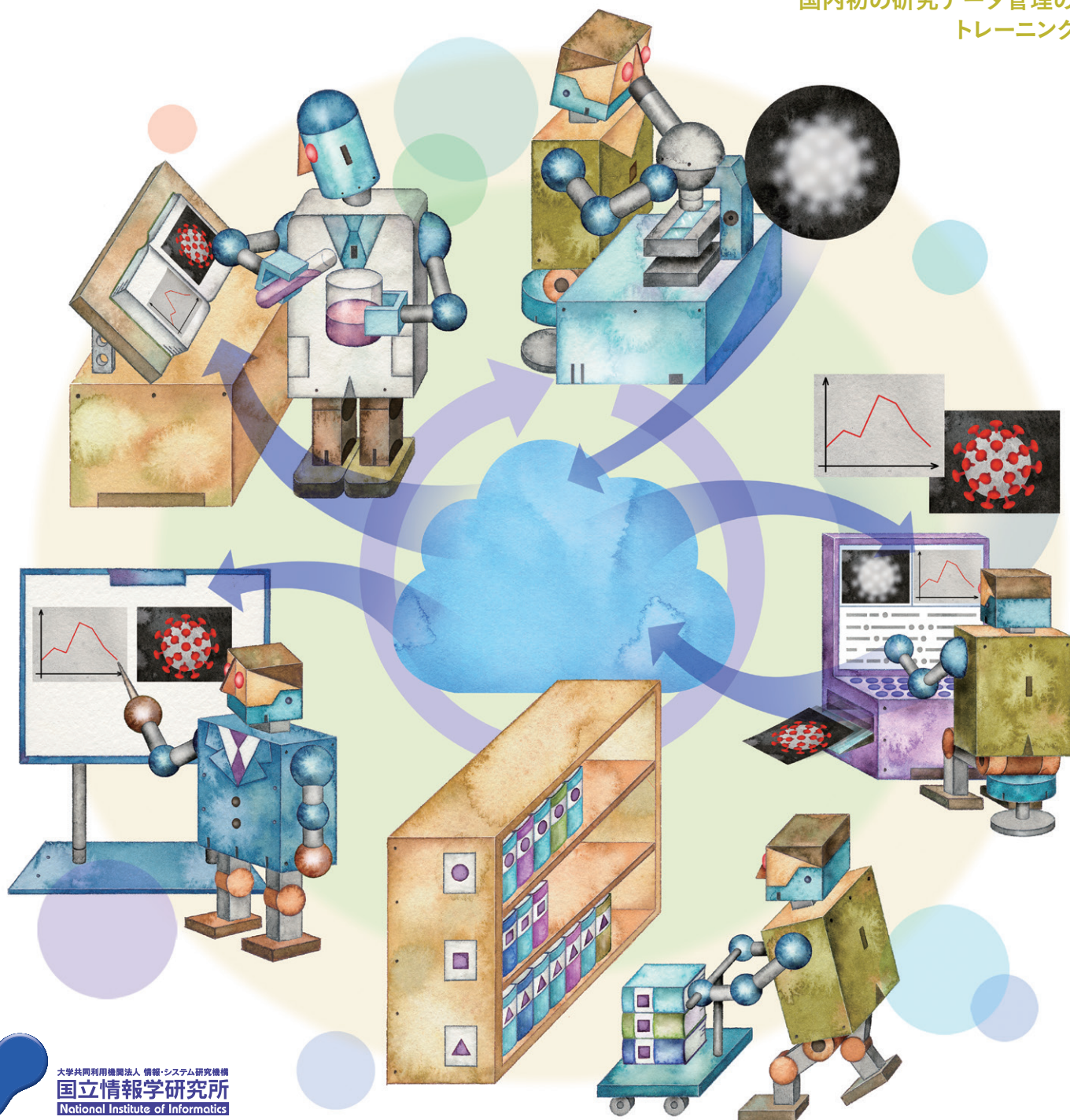
「WEKO3」

知の検索基盤として新たな役割を担う

「CiNii Research」

国内初の研究データ管理のための

トレーニングコース



次世代研究データ基盤 「NII RDC」への期待

オープンサイエンスを推進し、イノベーションを促すために

橋爪 淳氏 Hashizume Atsushi

文部科学省研究振興局 参事官（情報担当）

聞き手：滝 順一氏 Taki Jun-ichi

日本経済新聞社 編集局編集委員

政府の「統合イノベーション戦略」や「科学技術基本計画」はオープンサイエンスの推進を強調する。研究成果を社会や科学界に還元するなかで、研究データを積極的に公開して分野横断的な研究を育みイノベーションを促す。研究の透明性を確保する狙いもある。国立情報学研究所は、大学などがもっている研究データを適切に蓄積・管理・利用する研究データ基盤の提供を始める。文部科学省研究振興局で情報分野を担当する橋爪淳参事官に、データ基盤整備の意義を聞いた。

データのオープン化とDXで研究環境の活性化を

—— まずオープンサイエンスがいま、加速する背景についてお聞かせください。

橋爪 研究成果やデータを共有、公開してさまざまな人の力を結集して科学を進めていくことが重要になっています。これまでに得られた成果を活用しながら効果的に研究活動を進めてい

くことができるからです。研究者個人のレベルにとどまらず、学術の領域を超えて異なる分野の研究者も気づきを得る。学際的、分野融合的な研究活動を促します。「第4の科学」という呼び方もされますが、観察や理論、シミュレーションといった従来の研究手法とは異なる、新しい「データ駆動型」のサイエンスの進展にも貢献すると思います。

またデータの公開は、職業的な研究者ではないが科学に強い関心を抱く多くの人々の力を結集することにもつながり、いわゆる「シチズン・サイエンス（市民科学）」の視点からも重要だと思っています。さらに、研究プロセスの透明化、研究の公正性の確保にも貢献していきます。

—— ここ数年、日本の研究力の低迷が問題視されていますが、研究力の向上につながるのでしょうか？

橋爪 研究力の低下にはさまざまな要因があり、オープンサイエンスだけで簡単に解決するとは思いません。ただオープン化とは車の両輪の関係にある「研究のデジタルトランスフォーメーション（DX）」を含めて考えれば、日本の研究環境を再活性化する大きな力になるのではないかと思います。

文部科学省の科学技術・学術政策研究所（NISTEP）が公表した分析のなかで、私が注目しているのは、新たな研究の芽となる可能性のある研究の割合が日本は他国に比べて少ない点です。日本の科学研究は既存の主流をなす研究分野に集中する傾向が強いのではないかと。科学の進展のためには、やはり新しいことへのトライアルが重要ではないかと思います。そのためにはいろいろなデータや成果を共有し、皆がアクセスしやすくする。異分野の人の参入をしやすくし、気づきの機会を増やすことが研究力の強化につながります。



橋爪 淳

1994年早稲田大学卒業、旧科学技術庁入庁。文部科学省において産学連携、原子力損害賠償、科学技術・学術に係る制度改革などを担当し、在カナダ日本大使館一等書記官、広島大学教授を経て、2019年4月より現職。

—— DX は政府の第 6 期科学技術・イノベーション基本計画でも重視されていますね。

橋爪 新型コロナウイルス流行以前から IT 技術の活用は重要な課題でしたが、コロナ禍で抜き差しならない課題になりました。政府全体ではデジタル庁を設けて社会の変革に取り組みます。文科省でも萩生田光一大臣を本部長とするデジタル化推進本部を設置し、研究・教育やスポーツ、文化の世界のデジタル化を進めています。

デジタル化といっても、既存の情報を単純に「0」「1」のデジタル情報に置き換えるのではなく、新しいスタイルの活動を生み出し研究や教育の環境をレベルアップすることが重要です。先ほど述べたデータ駆動型の新しい科学の振興はその 1 つです。一方で、大型の研究設備や研究室をリモート化、スマート化して研究者がポストコロナの時代を迎えた後も時間と場所の制約を超えてアクセスし、利用できるようにすることも重要です。

そうした研究環境をしっかりと整えていくことを通じて、研究者がそれぞれのライフスタイルに合った働き方ができる「働き方改革」に貢献する部分もあります。また、自動化で実験の検証を容易にしたり、実験をサポートする技術者の匠の技、たとえば細胞培養のテクニックなどを所属や場所に縛られることなく、さまざまな場所で再現したりできる可能性が広がります。

—— データ基盤整備では海外が先行していますか？

橋爪 欧州連合（EU）が「欧州オープン・サイエンス・クラウド」と名付けた大規模なデータ基盤をつくらうとしているなど、さまざまな動きがあります。また、2016 年の主要国首脳会議（G7）茨城・つくば科学技術大臣会合で、日本が提案してオープンサイエンスの推進を決議するなど各国で問題意識を共有しています。このような状況を踏まえ、日本としてもデータ基盤の整備を加速していく必要があると思います。

課題はリポジトリの連携、ルールづくり、人材育成など

—— NII が取り組む次世代データ基盤整備に期待するところは？

橋爪 地球環境分野など従来からデータの蓄積と共有が進んでいる分野もある一方で、データを蓄えるリポジトリをもたない分野もあります。全体をつなぐ基盤づくりとサポートが重要です。大学共同利用機関である NII が大学全体に向けたデータ基盤のサービスを始める、しかも既存の基盤と競合するのではなく全体のメタデータ連携による緩やかな統合をめざすというのは、日本全体のデータ基盤をつくる上で非常に重要な取り組みであり、大いに期待しています。文科省としても、NII とともにしっかり取り組みたい。

—— では逆に、データを駆使したオープンサイエンスを広げていく上での課題はどこにあると考えますか。

橋爪 先ほどの発言と裏腹の関係になりますが、まずデータを安心・安全に保管できるリポジトリの整備が重要です。閉じた

リポジトリではなく相互に連携できるものでなくてはなりません。そこは次世代の「学術情報ネットワーク（SINET）」整備のなかで NII と文科省が連携して実現に向けて取り組んでいきます。それが 1 つ目の課題です。

2 つ目は、基盤ができたら各分野の研究者が情報分野の研究者と協働して、データ活用による新しいサイエンスをいろいろな分野で広めていくことが必要です。

ルールづくりも大切です。研究データをどう管理していくのか。オープンにすべきものなのか、関係者だけで共有すべきものなのか、オープン・クローズの戦略も考えねばなりません。さらに、個人データも含め社会活動のなかで生み出されるデータを研究で活用する機会が増えています。そうしたデータを扱うにあたってのルールを考える必要もあります。

データをマネジメントする人材の育成も重要になります。データサイエンスを各研究機関がばらばらに進めるのではなく、グッドプラクティスの共有、困った時のサポートなど、全体を横につなぐ活動もどこかに担ってもら必要があります。

その点で NII には強みがあります。NII は、データ基盤を担う中核的な機能とともに研究力をしっかり備えている。SINET など研究基盤のサービス提供と先端的なデータサイエンスの研究を両輪で展開していることから、研究現場のニーズを研究にフィードバックして生かすことができる。研究基盤の上にさまざまなサービスを展開していただき、こんなこともできるといった先端的な事例をつくり新しい時代の可能性を示してほしい。

—— デジタルへの移行は、取り組めばその先に良いことがあるわけですが、研究者にとって最初は面倒なことも多いと思います。現場に向けてはどんなメッセージを。

橋爪 「ピンチをバネにして新しい挑戦を」などと言われますが、課題が山ほどあるなかで新しい状況に対応することは容易ではありません。デジタル化やオープンサイエンス化を進めた先にどんな姿形があるのかを示したいですね。ルールづくりも規制というより、こうすればうまくできるという共有知にして、少しでも現場で悩まなくてもすむようにしていきたいと思っています。

インタビューからのひとこと

デジタル・トランスフォーメーション（DX）やデジタル化という言葉を見聞きしない日はない。新型コロナパンデミックを通じて、日本社会や政府のデジタル化の遅れを痛感させられた。デジタル化のすべてに対し諸手を挙げて歓迎はできないものの、否応なく潮流に乗り世界の流れに先回りする覚悟が今の日本には必要だと思える。

橋爪さんが指摘するように、日本の研究機関のなかで NII はユニークな立ち位置にある。SINET に代表される研究情報基盤の提供と情報科学の基礎的な研究力を併せもつ。パンデミック下で NII が開催したオンライン教育に関する一連のサイバーシンポジウムは、その機能を発揮した端的な例だ。急きょオンライン教育への転換を迫られた大学を情報共有によって支援した。データ基盤に関しても同様の動きを期待したい。



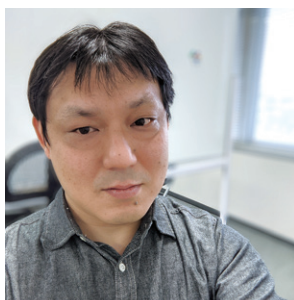
滝 順一

日本経済新聞社編集局編集委員
早稲田大学政治経済学部卒業後、
日本経済新聞社に入り地方支局
や企業取材を経て、1980年代半
ばから科学技術や環境分野を担
当してきた。著書に『エコうま
に乗れ!』（小学館）、共著に『感
染症列島』（日本経済新聞社）など。

写真＝佐藤裕介

NII×名古屋大学 NII 研究データ基盤「NII RDC」が いよいよ始まる！

見えてきた研究データ管理の課題と展望



青木学聡

Aoki Takaaki

名古屋大学 情報連携推進本部
情報戦略室 教授／
大学 ICT 推進協議会 (AXIES)
研究データマネジメント部会 主査



林 和宏

Hayashi Kazuhiro

名古屋大学附属図書館
情報管理課専門職員／
オープンアクセスリポジトリ推進協会
(JPCOAR) コンテンツ流通促進作業
部会 JAIRO Cloud チームメンバー



田中幸恵

Tanaka Sachie

名古屋大学附属図書館
東山地区図書課 東地区図書
統括グループ 係長



山地一禎

Yamaji Kazutsuna

国立情報学研究所
コンテンツ科学研究系 教授／
オープンサイエンス基盤研究
センター (RCOS) センター長

2021 年度から、NII 研究データ基盤 (NII Research Data Cloud) が本格的にスタートした。NII RDC は、研究データの管理基盤 (GakuNin RDM)、公開基盤 (JAIRO Cloud)、検索基盤 (CiNii Research) の 3 基盤から構成されており、今回、新たにサービスを開始する管理基盤については、2020 年に実証実験を行った。この中で、先駆的な取り組みをしてきた名古屋大学では、附属図書館や情報基盤センターが中心となって研究データ管理の環境整備を進めている。環境整備やルール策定、データ公開基盤の整備・支援などを行ってきた名古屋大学の各担当者を迎え、NII オープンサイエンス基盤研究センター (RCOS) の山地一禎センター長とともに、現在の取り組みや見えてきた課題、展望について語っていただいた。

研究データの管理に大学も危機感

—— 2021 年度より、NII Research Data Cloud (NII RDC) が本格的に稼働しました。従来のサービスとの違いや意義をお聞かせください。

山地 これまで NII は、デジタル化された学術論文や資料、研究成果などを保存・運用・公開する機関リポジトリを提供する「JAIRO Cloud」や文献検索サービス「CiNii」など、主に公開された論文に関するサービスを提供してきました。NII RDC では、こうしたサービスを進化させるとともに、学術研究活動の過程で生成された研究データや関連の資料などを保存・管理できるように、「GakuNin RDM (Research Data Management)」という研究データの管理基盤サービスを新規にスタートさせました (図 1)。これはまさに、研究活動に直結する新しいサービスであり、NII としても大きなチャレンジです。

オープンサイエンスの営みは、図 1 のようなライフサイクルで回ります。まず研究計画を立てるところからスタートして、実験などを通じてデータを取得し、解析を行って、論文などにまとめて成果を公開します。それが他の研究者の着想に結びつき、新たな研究へつながる。いままでと大きく違うのは、右側のクローズドの部分、つまりこれまで非公開だった研究中のデータを共用の基盤で管理する点です。これを公開基盤と連携させることで、研究者の判断によりオープンにしていける研究データは公開し、サイエンスの発展につなげていく狙いがあります。

これまで、大学では図書館が中心となって、機関リポジトリなどを通じ、論文の公開の部分を担当してきたわけですが、それを研究データの管理にまで拡張するということが、大学にとっても大きなチャレンジになりますね。

青木 研究データは研究者や研究室が個々の方法で管理しているのが実情で、まさにクローズドです。研究者の皆さんがどん

な研究データを扱っていて、それをどう管理しているのか、外からはわかりません。また、研究領域ごとに管理の仕方もさまざまです。研究データは研究者にとって財産ですが、講座制の廃止や特任制度の拡大で、教員の退職や異動とともに研究室のデータや研究ノート、ノウハウなどがそっくり失われてしまうような事態も起きています。このままでは大学としての財産を失うことになり、大学側は危機感を強めています。これは、アカデミア全体の損失でもあり、デジタル化によって効率的に保存・運用していくことが重要だと思います。

山地 ハードディスクが壊れて研究データが失われてしまったり、どこにデータを保存したかがわからなくなってしまうということは日本中で起こっています。研究者が個々にハードディスクやサービスを購入してデータを管理していたら手間やコストもかかります。また、現在、研究公正の観点から、研究資料などを原則10年保存することがガイドラインで定められていますが、十分に実行されているとはいえない状況です。大学が組織として研究資料やデータを管理できる環境を積極的に提供してくれれば、研究者にとってもメリットは大きいはずで、潜在的なニーズは非常に大きいと思います。

大学図書館の新たな役割

——名古屋大学ではどのような取り組みをされているのでしょうか？

青木 名古屋大学では2019年の初めに総長から直々に「研究データ管理やオープンサイエンスを推進するように」と指示があり、検討体制が急速に立ち上がりました。まず方向性を定めるために学術データポリシーを作成し、「研究データの管理・公開・利活用を支援するための環境を提供することが大学の使命であること」を明確にしました。現在は図書館や情報基盤関係の教職員が中心となって環境整備を進めています(図2)。

林 図書館では主に公開基盤に関する業務を行っています。公開部分は研究データのサイクルの中では、学術情報が循環し、次の新しい研究を牽引する重要な役割を担います。これまではNIIの機関リポジトリ・サービス「JAIRO Cloud」上に学術論文などを保存し、情報発信をしてきました。現在は、論文だけでなく、研究データも公開できる新しいJAIRO Cloud(ソフトウェア名はWEKO3)への移行作業を行っています。

山地 WEKO3は新しい機能を追加しやすいシステム構成になっています。現状はまだ使われ始めたばかりですが、これから各機関で運用が始まってフィードバックを得ることで、使いやすいように機能を拡充していく予定です。

林 研究データはこれまで扱ったことのない形式なので、図書館にとっては新たな挑戦になりますが、NII RDCを通じて、クラウド上で他の学術機関と課題やその解決方法を共有できるのは心強いと感じます。先行した事例が手本になって、次の事例が続くような連鎖が起これば、一気に広がりを見せて、日本中で学術情報の管理体制が整い、情報発信へつながっていくだろうと期待しています。

田中 研究中からデータを管理基盤に保存していれば、公開基盤への移行も簡単になります。そういったシステム連携のメリットの広報などが足りていないので、図書館では公開部分に限らず、研究データ管理の促進も含めて支援していけたらと思っています。研究データのオープン化の世界的な潮流に、図書館として寄与していきたいですね。

研究データ管理は誰がどう支援する？

——研究データ管理を実践するには、支援する人材も不可欠です。

田中 オープンアクセスリポジトリ推進協会(JPCOAR)の作業部会が中心となって、研究データ管理を支援する人材育成の

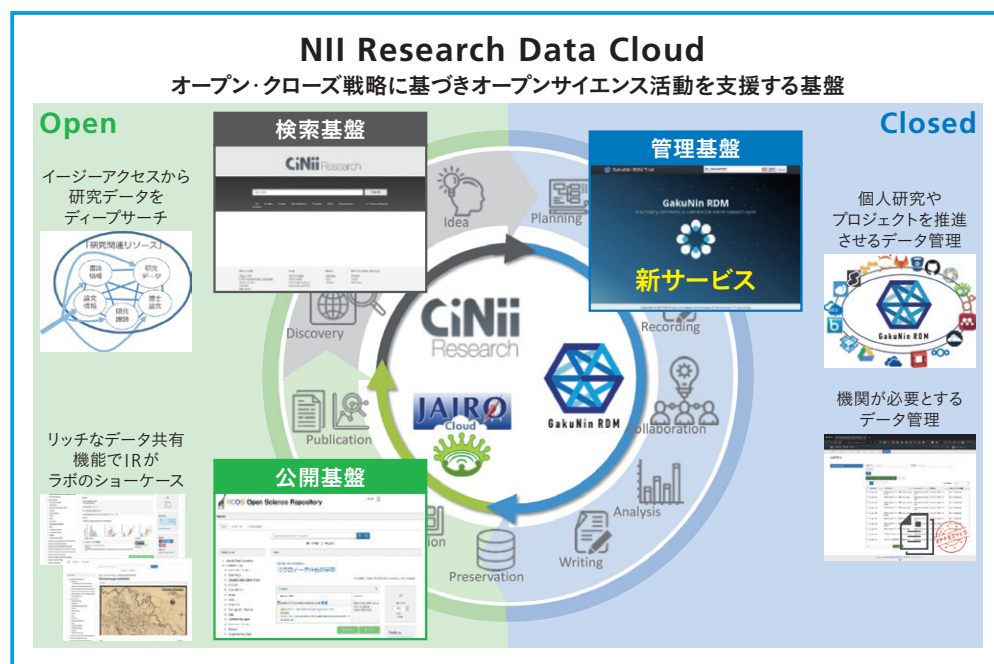


図1 NII Research Data Cloud

研究データのライフサイクルの中で、データ管理をGakuNin RDM、公開をJAIRO Cloud、検索をCiNii Researchがそれぞれ支援する。互いのデータ移行がスムーズで、ストレスな研究データ管理システムを提供する。

青木学聡

2020年4月から名古屋大学情報連携推進本部情報戦略室教授として大学全体の情報基盤の企画に携わる。前職の京都大学情報環境機構時代より、研究支援のための情報基盤整備の立場で研究データ管理の課題に取り組む。

林和宏

2018年に名古屋大学附属図書館に赴任し、リポジトリの運用・管理業務に関わる傍ら、JPCOARで、次期JAIRO Cloudの実証実験や移行機関サポートの活動に参加する。

田中幸恵

2008年4月より図書系職員として名古屋大学に勤務。2020年度学術情報ネットワーク運営・連携本部オープンサイエンス研究データ基盤作業部会トレーニングサブワーキンググループ委員として、研究データマネジメント推進にかかる人材育成のための継続的な教材作成・更新、運用体制についての検討を行っている。



図2 | 名古屋大学でのRDM支援体制の構築

名古屋大学での取り組みは、竹谷喜美江「名古屋大学における研究データ基盤整備推進組織の整備について」、第3回SPARC Japan セミナー2019 (<https://www.nii.ac.jp/sparc/event/2019/20200207.html>) も併せてご参照ください。

ための教材がいくつかつくられています。いまはそれを各研究機関でより活用してもらうため、NIIに設置されたオープンサイエンス研究データ基盤作業部会トレーニングサブワーキンググループ（以下TSWG）で、研究データ管理に関する一連の業務とその遂行のために必要なスキルをピックアップしたリストを作成しています。こうしたスキル一覧表を提示することによって、各研究機関において支援人材を育成するための教材やプログラムの選択・開発・評価に使っていただくことを意図しています。さらに、この業務を担い得るのはどのような職種の人なのか、どのような人材を育成すべきなのか、という情報も付け加えていますので、各研究機関の状況に合わせた組織的な支援・連携体制を構築する際のヒントにいただけたらと考えています。

山地 既存の教材は、もともと海外の事例や教材を参考にしてつくったものなので、TSWGではそれを日本の職制や大学の組織構造に合う形に再構築してくれています。

田中 図書館員全体を見ると、研究データ管理に興味はあるけれど、これまでの業務と直結するものではないので、どう関わったらいいかわからない人が多いという印象です。そうした人たちに向けて、一歩踏み出すきっかけになるようなコンテンツを提供していきたいと強く思っています。正直に言うと、私も当初、研究データ管理は本当に図書館が担うべき業務なのかと疑問に思っていました。ところが関連タスクを整理してみ

ると、まったく新しいことばかりではないことに気づきました。公開基盤に登録された研究データへのメタデータの付与など、いままでの図書館の仕事の拡張やプラスアルファでできそうなことが実はたくさんあります。

林 扱うデータの形式が多様になり、クラウド化されたという違いはありますが、図書館がこれまで行ってきた学術雑誌の整理や提供と基本的には同じ役割、つまり、学術情報の流通と保存によって、教育研究活動を支援していくという役割に変わりはないのです。私もそのことに気づいてからは、利用者に対してベストな提供方法を考えるようになり、やりがいや面白さを感じるようになりました。

青木 図書館や情報基盤センターが行っていることは、研究者や学生には意外と知られていません。研究データ管理に関わることで、研究者の生活により密着したところにリーチし、お互いに信頼関係を築く良いきっかけになるかもしれません。

田中 青木先生のおっしゃるように、研究者や学生に図書館等のポテンシャルを知ってもらい、研究データを公開したいときに頼れる場所としてもらえたら理想です。大学事務組織においても、これまで、図書館の仕事は外から見えづらく、アピールは難しいと感じていました。研究データ管理に関わる業務は多岐に渡っており、その支援を一部署で完結することは不可能です。どうしても、図書館以外の部署との連携が必要になり、お互いができることを共有しながら、組織的に支援を提供してい

くことになります。研究データ支援に関わることで、図書館のプレゼンスの向上にもつながっていただければ幸いです。

まずはニーズを掘り起こす

—— 2021年2月15日から管理基盤である GakuNin RDM の本格運用が始まりました。利用状況はいかがですか。

青木 名古屋大学では昨年の実証実験から参加していて、大学が用意したストレージを使って、それを GakuNin RDM に接続する形で運用しています。つまり、データ本体は名古屋大学のシステム内にあり、それを共有するための入口が GakuNin RDM になります。今後もこのスタイルで運用していく予定です。大規模な研究プロジェクトでの活用事例はまだありませんが、研究室の中で使っているという声はいくつか聞いています。

これまでの利用例を見てみると、GakuNin RDM の魅力は、研究室内や大学内での研究データ管理だけでなく、他の学術機関とのコラボレーションを進める際のポータルとしての使い方にあるのではないかと感じています。名古屋大学の事例ではありませんが、実際に大学間での共同研究に GakuNin RDM をポータルとして研究データを管理する体制をつくったという話もあります。

—— どのような課題が見えてきましたか。

青木 研究者の多くはすでに独自の方法で研究データを管理しているので、そこに新しい仕組みを浸透させるのは容易ではありません。「新しい酒は新しい革袋に盛れ」という言葉があるように、まずはこれから研究を始める若手研究者へのリテラシー教育や、新たに始まる研究プロジェクトに合わせて利用を勧めていくことが、効率的な利用拡大につながるのではないかと考えます。また、こうした新しい仕組みを普及させていくには、不慣れな環境を乗り越えるための支援体制を充実させていくことも不可欠です。

山地 新しいことをやるときは、必ず反対する人がいて、大きな課題に直面するものです。人の考えや行動パターンを変えるのはすごく難しいことですが、ニーズをきちんと掘り起こし、徐々にいまのワークフローを変えていくのは、実はいちばん面白いところでもあります。

青木 まさに名古屋大学では、ニーズを掘り起こしながら仕組みをつくっているところです。ニーズは確かに存在し、研究者にとって便利になるのは間違いありません。研究データの管理を独自にやろうとすると、データにどんな名前を付けてどこに保存するか、研究ノートにはどんな項目を書くかなど、決めなくてはいけないことが山ほどあります。GakuNin RDM の使い方が浸透していけば、同時に機能拡張も進み、データ管理の方法も標準化されていくはずですよ。それにより、研究者のストレスは格段に減るでしょう。

山地 最近、退職する教授が学術の発展のためにこの情報をぜひとも残しておきたいという場合や、論文発表の際に研究データも一緒に公開しなくてはならないケースが増えています。公開部分に関しても、そういうニーズをきちんとキャッチして、

徐々にすそ野を広げてワークフローを変えていく。将来を見越し、海外の動向に目を向ければ、それはやはり研究者にとって必要不可欠な営みになっていくと思います。

大学の声を聞きながら機能を実装できる強み

—— 今後の普及に向けて何が必要でしょうか。

青木 研究データ管理に関しては、単にストレージを共有するだけでは、他のさまざまなクラウドサービスと競合してしまいます。ストレージの共有から一歩先に進んだ、研究プロジェクトのポータルとしての機能を増やしていくことが必要ではないかと思っています。

山地 日々使うものだから研究者の目は厳しい。「こんなもの使えない」といわれないように、使いやすく、役立つものをつくらなくてはけません。研究データ管理にどんなスキルが必要かといった情報や、大学での活用事例からどんな機能が必要かといったフィードバックがこれから得られると思うので、NII は大学とともに PDCA を回しながら改善改良を重ねていく。それをこれからスタートしていくわけで、さらに重要なフェーズに入っていきます。

青木 いままで以上に協力体制を強化していく必要があります。機能としては、まずは手軽に使えることがもっとも重要だと思います。事前準備としてやるのがたくさんあると、使う前に気持ちが萎えてしまいます。GakuNin RDM のシステムは、大学が「学術認証フェデレーション（学認：GakuNin）」に参加していればアイデンティティプロバイダ（IdP）連携のみですぐに使えるようになります。ただ、IdP の運用は機関の問題なので、末端の研究者から見ると、非常にハードルが高くなってしまっているのが現状です。より多くの機関が学認に参加し、GakuNin RDM を気軽に利用できる仲間がどんどん増えてほしいと思っています。

田中 研究データ管理を支援する側にとっても、簡単に使えるものでないと、結局、指導できる人も限られて内々のものになってしまいます。NII が構築してくださったシステムをただ使うだけではなく、大学側が積極的にユーザーボイスや活用事例をフィードバックして一緒に良いものをつくっていくという姿勢で臨みたいと思います。

林 公開基盤に関しては、まだ日本国内での事例が少ないので、NII RDC が各機関と事例を共有できる場所、仲間を増やす場所として機能していけば、さらに発展していく足がかりになると思います。

山地 大学の声を聞きながら、必要な機能を実装できるのは、私たちの最大の強みです。大学との信頼関係は NII にとっての財産ですからね。今回、NII RDC を構築していく中で、大学の人たちや若い人たちとも同じ目線で会話ができるような体制が整ってきました。セクターを超えて一緒に大きな挑戦をしていくことは、若い人たちにとって非常に良い経験になるでしょう。これが成功すれば、いままでよりさらに強固な関係を築いていけると楽しみにしています。（文＝秦千里）

研究データ管理基盤「GakuNin RDM」の本運用がスタート

先進的な研究データ管理を支援する

2021年2月15日、研究データの管理を行う大学や研究機関を支援するサービス「GakuNin RDM (Research Data Management)」の本格運用が始まった。これは、NIIがオープンサイエンスのために提供する研究データプラットフォーム「NII Research Data Cloud (NII RDC)」内で、初めて提供する研究データ管理基盤として、オープンサイエンス基盤研究センター(RCOS)が開発したものである。オープンサイエンス実践の鍵となるRDMの基盤として、どのような機能を備えているのだろうか。



込山悠介 Komiyama Yusuke

国立情報学研究所 コンテンツ科学研究系／
オープンサイエンス基盤研究センター 准教授



新妻 聡 Niitsuma Akira

国立情報学研究所 学術基盤推進部 学術コ
ンテンツ課 研究データ基盤整備チーム 係長

サービスビジョンは「研究推進」と「研究公正」

近年、大型研究での先進的なデータマネジメントの実施に加え、研究助成機関に提出する計画書にデータ管理の記載が義務付けられるなど、研究を進める上でデータを組織化・構造化して保存・管理することは避けて通れなくなってきた。研究データ管理業務を適切に行うために、その管理業務は研究者個人から大学や研究機関などの組織に移り始めている。

この学術界の変化を支えるため、NIIは、NII RDC内で初めて研究データ管理基盤サービス「GakuNin RDM」を開始した。主要な目的は研究データを管理・共有するITインフラの提供だが、開発者を担当した込山悠介准教授は、本サービスの機能について「研究者には気軽に使ってもらい、機関には手堅い研究データ管理環境を提供する基盤でなくてはならない」と話し、「研究推進」と「研究公正」をサービスビジョンに掲げる(図)。

まず、研究推進の観点においては、NIIが提供している認証フェデレーションGakuNinの参加機関であればすぐにサービスを受けることができる。また、これまで蓄積してきた研究データを取り込めるよう、外部のクラウドやリポジトリとの連携を可能にした。データ共有など共同研究で必要となる機能も備えている。これらは研究者個人や機関など利用者の使いやすさに配慮した機能である。一方で、研究データの研究証跡をシステムが常に記録しており、研究不正や疑念行為が発生した際

には研究データの追跡調査ができる。この機能には研究不正の抑止効果もあり、機関の手堅いデータ管理を支援する。

安全なデータの共有と産学連携の支援をめざす

現在、利用できる機能については、2019年4月から2020年10月までの実証実験で、参加22機関がトライアルを行った。東京大学定量生命科学研究所は研究公正機能を自機関で活用する方法を検討し、北海道大学はデータ管理基盤と実験装置・計測装置との連結を計画するなど、機関ごとに独自のRDMの取り組みを行った。こうしたGakuNin RDMの利用実態を踏まえて、研究者が発表する論文、画像、生データなどをすべて適切に管理する仕組みや、実験装置から取り込んだデータを分析するプラットフォームとの連携、学内向けに提供されるストレージのRDM用途での活用など、幅広いニーズがあるとわかった。「なかでも多く聞かれたのが、NII RDCのデータ公開基盤であるJAIRO Cloudと接続してほしいという声と、産学連携のために企業のアクセスを可能にしてほしいという要望でした」と込山准教授。多くの機関が、GakuNin RDMを単なるストレージではなく、将来はデータ利活用や大規模な共同研究で使いたいと考えていた。なお、JAIRO Cloudとは2021年度中に接続する予定で、これによりスムーズなデータ公開が可能になるだけでなく、GakuNin RDMからJAIRO Cloudへのデータ受け渡しの手続きを経ることで、データの誤公開を防ぐことができる安全な研究データ基盤が構築される。

GakuNin RDMは2月15日に本格運用がスタートした。これを機に、サービスの運用業務が学術コンテンツ課に移った。組織的、法的な準備を担当した新妻聡係長は、「学術コンテンツ課では参加機関のうちの約10機関で構成する研究データ基盤運営委員会を立ち上げました。これからは運営委員会と議論しながら、RCOSとともに安定運用や機能強化、改善を続けます」と話す。将来的には、サービスの継続のために有料化も視野に入れているが、当面は無料での提供を予定している。日本の研究を活性化させる力を秘めたGakuNin RDMは、盤石の体制の下で充実したサービスを提供していく。(取材・文=池田亜希子 写真=古末拓也)

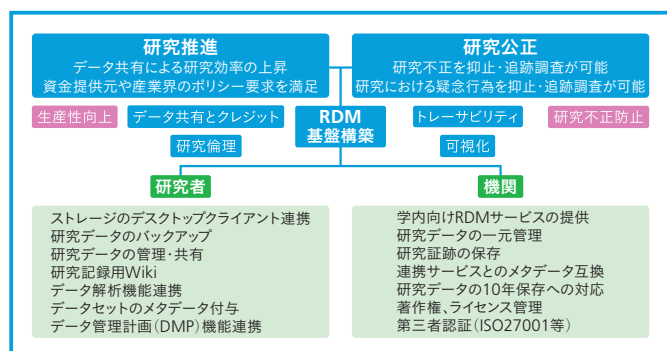


図 | GakuNin RDMサービスのビジョン。「研究推進」によって研究者を支援し、「研究公正」によって機関に手堅いデータ管理を提供する

日本の学術機関のデータ公開を支える「WEKO3」

JAIRO Cloudの基盤ソフトウェアWEKO3が始動

実は、日本は米国に次ぐ「機関リポジトリ先進国」であり、機関リポジトリをもつ大学や学術機関は876（2021年1月時点）に上る。その約7割が利用しているのが、NIIとオープンアクセスリポジトリ推進協会（JPCOAR）で共同運営する機関リポジトリサービス「JAIRO Cloud」だ。現在、JAIRO Cloudは、その基盤ソフトウェアである「WEKO」とともに次期バージョンへの移行・開発が進められている。



林 正治 Hayashi Masaharu

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 特任助教



林 豊 Hayashi Yutaka

国立情報学研究所 学術基盤推進部 学術コ
ンテンツ課 研究データ基盤整備チーム係長

JAIRO Cloudを研究データリポジトリとして刷新

機関リポジトリとは、デジタル化された論文や学術資料、研究成果などを大学や学術機関が保存・管理・公開するための一連のサービスの総称である。日本の機関リポジトリ導入数は世界でもトップレベルだが、自前でシステムを構築・運用できる学術機関は決して多くない。NIIはそうした機関を支援するため、2012年度から機関リポジトリ機能を提供するマネージドサービス「JAIRO Cloud」を提供してきた。初年度の利用機関は73だったが、いまや638機関がこれを利用している（2021年1月現在）。なお、サービスの運用は、大学などが会員として参加する「オープンアクセスリポジトリ推進協会（JPCOAR）」と共同で行っている。

サービスの運用を担当する林豊係長は、「JAIRO Cloudはサービス開始から、多くの利用機関や図書館コミュニティの意見や要望を取り入れながら機能の拡張をしてきました。いわばコミュニティとともに成長してきたクラウドサービスです」と語る。

大学や学術機関といったエンドユーザーのニーズに継続的に応えていくことに加え、JAIRO Cloudにはもう1つ、学術界から期待される役割がある。ここ数年、文献資料だけでなく、研究データへのオープンアクセスが世界的に求められるようになってきたのだ。こうしたトレンドを受けて、NIIは、JAIRO Cloudの基盤ソフトウェア「WEKO2」の機能性を強化し、クラウドサービスとしての可用性や信頼性、保守性を担保しながら、文献リポジトリと研究データリポジトリとしての機能要求に応えていくことを決断。先端的なアーキテクチャを数多く採用した「WEKO3」を基盤とする新たな機関リポジトリサービスとして、JAIRO Cloudを刷新することにした。

WEKO3の開発に携わる林正治特任助教は、「WEKO2は利用者の声に寄り添った、きめ細やかな文献リポジトリサービスとして進化してきました。その良さは継承しながら、WEKO3では研究データリポジトリとしての役割をも担う新たな学術情報の流通基盤として構築し運用していきます」と意気込みを語る。

WEKO3はWEKO2をほぼ全面的に書き換え、インフラ構成もこれまでとは大きく異なる。パブリッククラウド上にKubernetesクラスタを構築し、マスターデータベースにはPostgreSQL、メッセージブローカーにRabbitMQ、全文検索にElasticsearchなどを採用、パフォーマンスや負荷分散、運用の自動化などに配慮したモダンなアーキテクチャとなっている。これは欧州原子核研究機構（CERN）が開発したリポジトリソフトウェア「Invenio 3」を下敷きにしており、林特任助教はその理由を、「細やかな機能の実現に最適で、かつ学術世界の潮流にも沿ったモダンなフレームワークだから」と説明する。

これにより、高い自由度でメタデータを扱えるようになり、一般の閲覧者、登録ユーザ、リポジトリ管理者の誰が利用しても閲覧や検索、カスタマイズなどがしやすいサービスが誕生することになった。

開発と運用の両輪で継続的な発展をめざす

WEKO3をベースにした次期JAIRO Cloudへの移行は600以上の利用機関に影響するため、2021年3月から段階的に実施している。移行ドキュメントの作成、連携サービスとの調整、データ変換と利用機関による確認、サービスの切り替えなど、やるべき作業は山積している。重要なのは、開発と運用のチームが同じ方向を向きながら、ユーザーサポートを担うJPCOARの関係者も含めて、密に連携していくことにある。

「開発と運用は完全に切り分けられません。お互いに丁寧なコミュニケーションを心がけ、双方の視点から細かな改善を継続的に取り入れていくことが重要です」と林係長は言う。両者が見ているゴールは「研究成果の活用が進む学術社会をつくること」であり、WEKO3/JAIRO Cloudは、間違いなくその礎となるだろう。

「アジャイル開発にも取り組んでいるが、試行錯誤の段階でうまくいかないことも多い。それでも挑戦をやめることなく、迅速な研究開発と安定したサービス運用の両立をめざしていきたい」と林特任助教。日本の機関リポジトリサービスを支えるための試行錯誤は続く。（取材・文＝五味明子 写真＝佐藤祐介）

知の検索基盤として新たな役割を担う「CiNii Research」 サイテーションからリレーションへ、オープンサイエンスへの道を拓く

NII が提供する学術検索基盤「CiNii」は、国内の論文を探す「CiNii Articles」、全国の大学図書館に収蔵されている書誌を探す「CiNii Books」、日本の博士論文を探す「CiNii Dissertations」という3つのサービスから成る。現在、これらを統合し、加えて論文にひもづく研究データの検索も可能にする新たな検索基盤「CiNii Research」の開発が進んでいる。その先行サービスとして「CiNii Research プレビュー版（以降、プレ版）」が2020年11月から公開されていた。



大波純一 Onami Jun-ichi

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 特任准教授



阪口幸治 Sakaguchi Koji

国立情報学研究所 学術基盤推進部 学術コ
ンテンツ課 学術コンテンツ整備チーム係長

3つのサービスをまとめて一律の検索を可能に

CiNii Research が CiNii と大きく異なる点は3つに分かれていた検索サービスが統合されること、研究データや研究プロジェクト情報も検索対象としていることである（図）。インフラには CiNii の技術はほとんど踏襲せずに、新規につくり上げた。検索基盤はパブリッククラウド上で構築され、外部のデータベースや海外の学術機関とも API 連携されており、さらに検索エンジンにはオープンソースの「Elasticsearch」を実装するなど、モダンでパフォーマンスにすぐれた構成だ。

大幅な更改に至った背景について、開発リーダーの大波純一准教授は「論文と研究データや関連資料の関係性をたどり、ある研究データを別の研究に再利用する機会を増やすという、サイテーション (citation) からリレーション (relation) へのニーズがサイエンスの世界で急増しているため」と説明する。近年、学術の世界ではオープンサイエンスの考え方が浸透しており、それに伴って研究データも公開する傾向が強くなっている。また、既存の CiNii はリソースごとに特化した枠組みとなっているため、統合されたインターフェースから横断的に検索を行いたいと望む声も大きくなっていた。こうした時代の変化と要請に応えるべく、NII は2017年から CiNii Research の開発を開始、2021年4月の正式公開に向けて着々と準備を進めてきた。

関連の研究データやプロジェクトも検索可能に

機能を進化させるために、CiNii Research ではオリジナルなグラフデータベースを構築している。グラフデータベースは、個々の情報（論文、研究データ、書誌など）のエンティティ（実体）をグラフにおけるノード（節点や頂点）として、それぞれのエンティティ同士の関係性をエッジ（線）を使って表現する。1つのエンティティにつながるエンティティの数や種類には制限がないためデータの追加が容易なうえ、関係性をより速く深く検索することが可能だ。初期からプロジェクトに関わってきた阪口幸治係長は、「これまでは一度検索すれば、その時点で終わりましたが、1つのエンティティにつながるほかのエンティティ

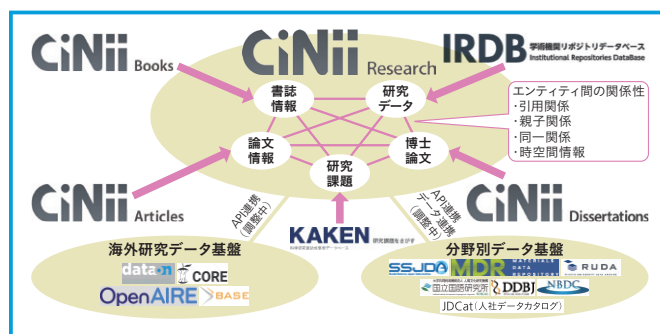


図 | CiNii Research がめざす適用範囲

を次々と横断的に検索できれば、オープンサイエンスがより加速していくでしょう」と期待する。

もっとも既存のサービスはそれぞれが異なるスキーマ（構造）を使ってメタデータを格納していたため、統合にあたり、「新しい基盤ではいかにメタデータを格納すべきか、その標準化にかなり頭を悩ませました」と大波准教授。最終的には JPCOAR が策定している『JPCOAR スキーマ』をベースにした。JPCOAR は公開基盤「WEKO3/JAIRO Cloud」の運用パートナーでもあり、基盤連携のしやすさの面でもメリットがある。

公開されていた CiNii Research プレ版について、利用者からは「こういう検索基盤を求めている」「検索スピードが速くなった」「1つの検索基盤から利用できるのは便利」などかなり高い評価が寄せられた。今後はフィードバックを取り入れながら改善を続け公開後も随時アップデートを重ねていく。

阪口係長は「CiNii Research は新規に開発したが、検索のしやすさや表示の見やすさなどは引き継がれています。また、日本語の検索が可能な基盤は英語などに比べて難易度がかなり高いのですが、CiNii ではその最適化も図ってきました」と語る。日本語による潤沢な学術情報を見やすい形式で検索できるという CiNii の使命は、そのまま継承されているというわけだ。

大波准教授は CiNii Research を「現代の図書館員の発展型」と表現する。日本の学術資産を継承しつつ、変化の潮流を受け入れた CiNii Research は、日本の知の検索基盤として新たなスタートを切った。（取材・文=五味明子 写真=佐藤祐介）

国内初の研究データ管理のためのトレーニングコース

体制づくりから支援人材の育成まで

2021年6月、NIIが全国の大学等と構築する学術認証フェデレーションの学習管理システム「GakuNin LMS (Learning Management System)」において、研究データ管理を始める機関向けのトレーニングコースの提供が正式に始まる。本コースをオープンアクセスリポジトリ推進協会(JPCOAR)とともに開発してきた2人に、日本における研究データ管理の現状と、トレーニングコースを実践へつなげていくポイントを聞いた。



古川雅子 Furukawa Masako

国立情報学研究所 情報社会相関研究系 助教／オープンサイエンス 基盤研究センター



南山泰之 Minamiyama Yasuyuki

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 特任技術専門員

研究データ管理の意義を知り自機関に適した体制を築く

トレーニングコースで提供される教材の開発は、研究データ管理に関わる人材育成を目的に、機関リポジトリによる研究データの公開・流通をめざす大学図書館等で組織されたJPCOARによって、2015年頃から始まっていた。その内容は、欧米の研究データ管理の事例を参考にしつつも、日本の大学や研究機関の事情に合わせて再構成されている。NIIはJPCOARと協働し、教材をMOOCや学習管理システム「GakuNin LMS」で試験的に提供してきた。開発に携わった南山泰之特任技術専門員は、開発のポイントを次のように語る。「大学における研究データ管理については、データの公開を見越して、情報の組織化、論文公開の経験を培ってきた図書館の主導が期待されます。ただし、図書館の立ち位置は国や機関によって違うため、各々の事情に即した実践的な研究データ管理体制を築かなくてはなりません。日本では、これまで図書館職員が研究データ管理に直接関わることがなかったため、URAや情報基盤センターとの連携が欠かせないのです。この点を意識した教材づくりをしています」

トレーニングコースは、研究データ管理の必要性を確認することから始まる。その理由を、教材開発からLMSの構築までを担当する古川雅子助教は次のように説明する。「これまで研究者が独自に行ってきた研究データの管理を、大学や研究所が組織

として継続的に支援することによりオープンサイエンスを促進し、研究データの分野横断的な活用を促すなど新たな科学的知見を得ることにまでつなげていくことができます。その基礎として研究データ管理が重要であると理解してもらうためです」

コース内容をすべて視聴すると、研究データのライフサイクル(生成→加工→分析→保存→公開→再利用)で求められるサポート(サービス)の全体像が把握でき(図)、自機関で研究データ管理を行う際に守るべきポリシーや、協力を仰ぐべき部署、育成すべき人材やそのスキルが明らかになる。これに基づき、それぞれの機関が研究データの管理体制を構築して実践することになる。

よりよいコース提供のため継続的に開発

トレーニングコースはMOOCを含めると、すでに数千人が受講し、その多くが、「研究データ管理の概要を系統立てて理解できた」と評価している。ただし、機関ごとに必要な教育内容が異なることや、短い時間で効率よく学習したいという意見が聞かれたため、コース内容を数分の単元ごとに切り分けたマイクロコンテンツとして提供することにした。これにより、各機関がコンテンツを自在に組み合わせてカスタマイズできるようになった。その結果、「学習者が学習しやすい環境を提供するだけでなく、教材の改訂がしやすくなった上に、細かな学習履歴データの解析により学習者が難しいと感じている内容を把握しやすくなりました」と古川助教。こうして、最新かつ学習効果の高いトレーニングコースを継続的に提供できる仕組みが整った。

今後は、研究者向けの教材をトレーニングコースとして提供する予定だ。南山特任技術専門員は、「研究データ管理を、論文発表や助成金をもらうための義務ではなく、研究活動における利便性や発展性に寄与する営みとして定着させたい」と語る。古川助教も「研究データ管理は、研究リテラシーとして学生を含め研究に携わるすべての人が学ぶことが望ましい」と話す。社会の求めに応じた教材開発は続くが、まずは2021年6月から正式公開されるトレーニングコースが、国内機関の研究データ管理をとれだけ後押しできるかに注目したい。

(取材・文＝池田亜希子 写真＝古末拓也)

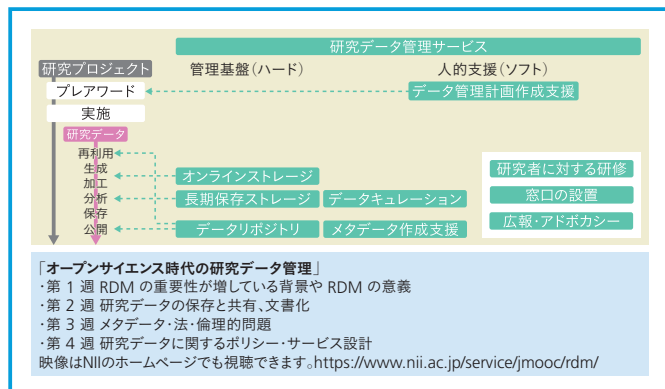


図 | 研究データ管理サービスの全体像(上)とトレーニングコース「オープンサイエンス時代の研究データ管理」の内容(下)。トレーニングコースURL: <https://lms.nii.ac.jp/>

大学間ネットワーク のあけぼの

安達 淳

Adachi Jun

国立情報学研究所 副所長

NIIの20年史「国立情報学研究所二十年の歩み」をまとめるなかで、SINETこそがNIIから大学への最大の貢献である、という思いを抱きました。そのなかでぜひ書き残しておきたいのが、大学間ネットワークの黎明期のことです。

1970年代にすでに現在の学術情報基盤の萌芽といえる構想がありました。その実現のために科研費により大規模な共同研究が次々と行われました。初代「お茶の水博士」の島内武彦東京大学大型計算機センター長が代表を務め、北川敏夫先生などがリーダーシップをとった特定研究「広域大量情報の高次処理」(1973年～)をきっかけに、「学術情報システム」構想が生まれました。1976年からの特定研究「情報システムの形成過程と学術情報

の組織化」では島内先生、次いで猪瀬博東大大型計算機センター長(後の初代NII所長)が代表を務め、ここで浅野正一郎先生を中心にコンピュータ間を接続する方式としてN-1プロトコルが開発されました。

これが実際に稼働し始めたのは1979年で、当時の東大大型計算機センターの鷹野澄先生が運用開発の中心となり、筆者も大学院生として接続実験の下働きを少し手伝いました。この通信方式はARPANETのTCP/IPに対し、CCITTが標準として定めたパケット交換X.25をネットワーク下位三層として使うものでした。ところが、7大学の大型計算機センターをつなぎはしたものの、通信料金をどのようにまかなうかという頭の痛い問題が出てきたのです。1980年から始まった日本電信電話公社のDDXパケット交換サービスでは1パケットが1円もして、これを大型計算機センターと利用者が折半するという暫定料金案で始めたのですが、利用が拡大する気配はありませんでした。

1984年に大学図書館を接続して図書の目録システムを動かし始めても、通信料金は悩ましい問題でした。翌年、文部省の学術情報課長に大変元気の良い西尾理弘氏が着任し、大学共同利用機関の学術情報センターを創設す

るという話をぶち上げました。筆者らは西尾氏の勢いに押されるようにしてセンター創設の概算要求を書き始めました。その議論のなかで当時助手だった橋爪宏達先生が「NTTがサービスを開始した光ファイバーの高速デジタル回線とX.25交換機を全部買ってしまおう」というアイデアを出し、とんとん拍子で実現したのです。これが従量制料金などの心配がいらぬ自営網のSINETのルーツです。

筆者が米国に滞在していた1986年に元Bell研のJohn Pierce先生にお目にかかった折、「日本ではT1ではなくX.25で専用網を作り始めた」と自慢したら「あんなややこしいやり方でスピードがでるのかな?」と言われました。どうやら256kbpsくらいがやっとだと後で知って少しがっかりしました。その後、この学術情報ネットワークはCSNETとBITNETの国際電子メールに接続したほか、JUNETへの回線の提供、ATMの導入と凋落など、さまざまな紆余曲折を経ました。しかし21世紀に入ってNIIに引き継がれ、高速性と信頼性を高いレベルで具備したSINETとして確実な地歩を築くことができました。

改めて、この文章で使った通信用語のほとんどがすでに忘れ去られたことを思うと、万感の思いです。

今後の予定

6月14日～19日 | Japan Open Science Summit 2021＝オンライン開催。詳細は、以下のURLで。
<https://joss.rcos.nii.ac.jp/>

6月18日～19日 | 国立情報学研究所 オープンハウス2021
(研究成果発表・一般公開)

6月末～7月上旬(未定) | 国立情報学研究所 学術情報基盤オープンフォーラム2021

※オープンハウス、オープンフォーラムとも、詳細が決まり次第、NIIウェブサイトでお知らせします。<https://www.nii.ac.jp/>

表紙の言葉

ロボットたちはそれぞれ研究、研究データの管理、公開、再利用を行っています。このオープンサイエンスのサイクルのなかで科学的知見が得られ、それが新たな研究分野につながると期待されます。NII研究データ基盤「NII RDC」は、この営みを強力に支援します。

情報から知を紡ぎだす。

NII

国立情報学研究所ニュース「NII Today」第91号 令和3年3月

発行 | 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所
〒101-8430 東京都千代田区一ツ橋2丁目1番2号 学術総合センター

発行人 | 喜連川 優 編集長 | 佐藤 一郎

表紙画 | 城谷俊也 編集 | 田井中麻都佳

制作 | 株式会社マツダオフィス / サイテック・コミュニケーションズ

本誌についてのお問い合わせ | 総務部企画課 広報チーム

TEL | 03-4212-2028 FAX | 03-4212-2150 e-mail | kouhou@nii.ac.jp

「NII Today」で
検索!



情報犬ビット
(NIIキャラクター)

<https://www.nii.ac.jp/today/>