

藤原 一毅

FUJIWARA, Ikki

国立情報学研究所 アーキテクチャ科学研究系 准教授

[研究紹介]

■研究者にはハードルが高いデータ処理環境の構築

現代の研究活動はすべて、データから叡智を見出す営みであると言ってよいでしょう。学術論文とその根拠データを原則公開し、データに基づく研究の発展を加速しようとする「オープンサイエンス」の方法論があらゆる学術分野に浸透しつつあります。NII では、オープンサイエンス基盤研究センター（RCOS）が中心となって、学術論文および研究データを管理・公開・検索できる研究データ基盤「NII RDC」を開発し、全国の大学や研究機関に提供しています。私は2018年に着任して以来、NII RDC のさまざまな拡張機能の実行基盤となる「データ解析基盤」の開発に取り組んでいます。

データはコンピュータによって処理されて初めて意味を持ちます。膨大なデータを研究者が逐一目視して意味を見出すことはできません。よって、公開データを活用しオープンサイエンスを実践するには、データとコンピュータの両方に精通しなければなりません。このうちデータについては、NII RDC をはじめとして、対象分野の専門家でなくても容易に発見・利用できるサービスが出揃いつつあります。一方コンピュータについては、データ処理に用いるソフトウェアを開発または入手し、処理の性質に適したハードウェアの利用権を調達し、そのハードウェア上で目的のソフトウェアが動作するよう適切な実行環境を構築する必要があります。この一連の作業は計算機システムの専門家にとってすら難しいことです。ましてコンピュータの専門家でない多くの研究者にとって、公開データを活用しオープンサイエンスを実践するうえで大きなハードルになっていると私は考えました。

■知識発見に専念できるデータ解析機能をすべての研究者に

私は手始めに、NII RDC のデータ管理基盤「GakuNin RDM」の拡張機能として「データ解析機能」を開発しました。研究者は、プロジェクト画面からボタンひと

つでデータ解析環境をクラウド上に展開し、GakuNin RDM で管理されているデータの解析プログラムをすぐ書き始めることができます。この機能により、研究者はプログラムの実行環境を自前で用意する煩わしい作業から解放されました。また、たとえば教員が構築した実行環境を多数の学生に複製させることも簡単です。研究者のニーズに対応して、大学や研究機関が所有する計算機をクラウドに代わって利用できる「計算機持ち込み機能」の開発も完了しています。さらに、データ解析機能を使って得られた知見をデータや解析プログラムとともに公開できる「計算再現パッケージ機能」も開発中です。これらの機能を GakuNin RDM とともに全国の大学に提供することで、データとコンピュータの両方に精通しなければならないという障壁を取り除き、さまざまな分野の研究者が公開データを活用しオープンサイエンスを実践する後押しをしてまいります。

データ駆動型科学への参入機会をすべての研究者に提供することが NII の使命であるというのが私個人の見解です。その手段として、GakuNin RDM を核としてデータの世界とコンピュータの世界を結び、データをコンピュータで活用する障壁を限りなく低くすることが私の目標です。専門性も習熟度もまちまちな、あらゆる学問分野の研究者や学生たちが、コンピュータの手配や管理に煩わされることなく、知識発見につながるデータ解析に専念できる環境を提供したいと考えています。

■私が思い描く GakuNin RDM を核とした次世代基盤サービス

私は 2024 年 1 月に RCOS からクラウド基盤研究開発センター（CCRD）に転属しました。NII RDC には今後、秘匿解析機能、データガバナンス機能、データプロビナンス機能、キュレーション機能、セキュア蓄積環境といった高度な機能が実装されていく予定です。私が手がけたデータ解析基盤は、これら高度化機能の実行基盤としても活用されていきます。また、CCRD が開発・提供している IoT (Internet of Things) 向けデータ収集基盤「SINETStream」と GakuNin RDM を連携させ、収集したデータの解析から解析結果の保存・共有・利活用までを統合的にサポートするサービスも開発していきます。

データ駆動型科学が普及していくにつれて、研究データはますます多様化・大容量化していくでしょう。個人情報や企業情報を含むセンシティブなデータも安心して扱える秘匿解析機能や、スーパーコンピュータで解析するようなビッグデータを GakuNin RDM で扱えるようにする仕組みなど、これからも研究者のニーズに沿った技術開発に取り組んでいきたいと考えています。

文部科学省が推進する研究データエコシステム構築事業には、私たち RCOS と CCRD が「研究データ基盤高度化チーム」の中心的メンバーとして参画しています。

私たちは、高度な基盤サービスの提供を通じて、データ活用による研究活動の変革（研究 DX）にも貢献できるものと期待しています。☐

Researchmap <https://researchmap.jp/ikki>

構成・執筆＝佐藤 尚規 2024 年 1 月更新