

政府関係機関事業紹介

研究データ時代の新しい大学情報システム

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. はじめに

国立情報学研究所（NII）が大学や研究機関に提供するサービスは、多岐にわたります。学術情報ネットワークSINETや、大学間の認証連携の仕組みを提供する「学認」、クラウド導入・活用支援サービス「学認クラウド」は、大学情報システムの高度化に携わる皆さまにも身近なサービスとして、ご活用頂いています。一方で、NII事業の中でも、学術コンテンツサービスは、論文や本のデータベースや検索サービスが主であるため、図書館職員以外の方々には馴染みが薄いかも知れません。

本稿で紹介する新しいサービスは、これまでの学術コンテンツサービスと密接に関係しながらも、大学情報システムにとってもミッションクリティカルなサービスとして位置づけられることになりそうです。これからの研究データ時代の基盤として、幅広い利用者から必要とされるサービスとなるように、現在、構築を進めています。今回から数号にわたり、その新しいサービス、NII Research Data Cloudの全容についてご紹介する予定ですが、本稿では、各サービスの詳細な説明に入る前に、前提となるオープンサイエンスと研究データ管理の関係や、NII Research Data Cloudを機関として導入する意義について解説します。

2. オープンサイエンスと研究データ管理

オープンサイエンスとは、論文だけでなく研究データを含む研究成果を積極的に公開することで、その再利用や分野横断型研究への発展を促進しようとする新しい研究のスタイルを指します。研究成果の再現性や透明性が確保されるため、研究公正の見地からも意義があります。この理念や可能性は理解できるのですが、実際に自分がやることを考えたら、どうしても消極的になります。公開するためには、研究データをどのように取得・解析したかを他の研究者にもわかるように記述する必要があります。実験条件を勘違いされて、

研究データを誤用されても困ります。公開した研究データを使って、他の研究者が新しい成果を得たら損をした気分になるかも知れません。オープンにしたい理由が先行してしまいます。

オープンサイエンスは、第5期科学技術基本計画（2016年1月）や統合イノベーション戦略（2018年6月）の中でも言及され、現代の重要な科学技術政策として位置づけられています。研究費の助成機関でも、オープンサイエンスを推進するための施策が準備されつつあります。オープンサイエンスに向けての政策的な動きは、着実に進んでいます。一方で、積極的に研究データを公開することにより、思いがけない研究の発展や産学連携に繋がったというボトムアップな事例の蓄積は、まだまだこれからです。政策的な動きとボトムアップな事例を融合させるためには、研究データを公開することへの負の印象を払しょくできるように、オープンサイエンスを体感した研究者をできるだけ増やす地道な努力が必要です。

研究データをオープンにすることに抵抗を感じる研究者も、研究室や学外の共同研究者との閉じた環境での研究データの共有は不可欠です。研究データを適切に保存し共有できる環境をもつことは、日々の研究を円滑に進めるための必須の要件になりつつあります。しかしながら、多くの研究者がその環境の整備に四苦八苦しているのではないのでしょうか？ 簡便に利用できるフリーミアムなクラウドストレージを使っているケースも多く見かけますが、セキュリティインシデントが発生した際に、適切に対処できるか不安です。既製のサービスに縛られて、本来必要とする機能が得られないこともあります。現状で望めるのは、限定されたメンバーでのファイルの共有機能ぐらいです。研究室や共同研究で必要とされるのは、そうした安易な共有ストレージではなく、研究者間で研究データや処理内容を構造的に引き継ぎ、適切に保存していけるシステムです。研究不正対応のためのガイドラインを守るためにも不可欠です。

皆さまとの情報交換を重ねると、研究データをオープンにする以前に、研究データの管理をサポートしてくれるサービスが必要とされていると実感します。逆に、研究データが適切に管理されていれば、研究データを公開するにあたって余計な労力を割かずに済むというメリットもあります。

3. 研究データ時代の機関の役割

ビッグデータやIoT、AIに関係した研究だけではなく、あらゆる分野の研究でデータを扱う機会が増えています。さらに、それを研究室内や共同研究者間で共有する機会も増え、研究の推進には、適切な研究データの管理と共有が不可欠になってきています。これは、個別の研究者だけの問題ではなく、研究者間の共通の課題です。すなわち、データの管理と共有ができる適切な環境をどう提供していくか、それを適切に実施する研究者をどう育成していくか、という課題が浮き彫りになってきます。

海外では、2010年頃にeScienceやデータ駆動型科学という言葉が使われだしたあたりから、大学の研究力強化のための重要な要素としての研究データ管理、そのための共有基盤の構築や人材育成への組織的な取り組みが始まっています。それを受けての、現在の国際的なオープンサイエンスの議論なのです。我々は、一足飛びにオープンサイエンスを目指すのではなく、こうした海外との状況の違いを知り、何が今の研究者に必要とされているかを考えるのが妥当です。

具体的なアクションを起こすための材料として、研究者が必要とする研究データの管理システムや、組織としてやるべきことが原則としてわかっていても、それを実現することは容易ではありません。限られた予算や人員のなかで、それぞれの組織が個別に検討を重ね、必要とされる新しいシステムを導入するには限界があります。研究プロジェクトで学外の共同研究者と研究データを共同で管理していくためには、組織

という単位ごとに環境が用意されても利便性に欠けます。データを管理し共有するための環境の構築は、組織間で競争しながら個別に取り組むべきではなく、組織間の協調領域として捉え、お互いに協力しながら実現していくのが効果的です。こうした背景に基づき、我々が新しく準備を進めているのが、研究データを管理・共有していくための基盤であるNII Research Data Cloudです。

4. NII Research Data Cloud

NII Research Data Cloudは、研究データの管理基盤、公開基盤、検索基盤の3つの基盤から構成されます。公開基盤と検索基盤は、それぞれNIIが既に提供している機関リポジトリのクラウドサービスJAIRO Cloudと文献検索サービスCiNiiを、研究データも扱えるように拡張することで実現します。管理基盤は、研究を実施する過程でデータを管理し、内部共有するための基盤で、NIIとして新しく提供するサービスです。公開基盤における共有と日々の研究の過程での共有では、必要となる機能が異なります。研究者のニーズを満足するためにサービスを分離しつつも、3つの基盤を有機的に繋げることで、オープンサイエンス時代の研究ワークフローを支える環境の構築を目指しています。次号以降では、それぞれの基盤の狙いと開発の状況についてお伝え致します。

