



デジタル時代の学術枠組み 「オープンサイエンス」 研究データの利活用をイノベーションに繋げる

船守美穂 Miho FUNAMORI

オープンサイエンスが花盛りである。2015年に、内閣府に設置された「国際的動向を踏まえたオープンサイエンスに関する検討会」が報告書を出したときは、その2年前のG8の共同宣言へのお約束を果たしている程度であったのに、今年夏に政府が発表した「統合イノベーション戦略」では、日本を「世界で最もイノベーションに適した国」に導くために掲げられた3つの「知の源泉」の1つに、「オープンサイエンスのためのデータ基盤の整備」が位置づけられた。ほかの2つの知の源泉も、データに関わりの深いものであるから、今回の戦略で政府がどれだけ「データ」というものを重視したかがわかる。では、オープンサイエンスとは何か、データとの関わりは何なのかを、ここで簡単に紹介したい。

オープンサイエンスとは

「オープンサイエンス」は、非常にわかりにくい用語である。学術は本来オープンなものであるという理解が一般にはあるから、何を今さら？ という印象をまず持たれる。ちなみに、学術が他者の知見の上に新たな知見を積むというオープンな営みとなったのは、1665年に英国王立協会から“Philosophical Transaction”が初の学術雑誌として創刊されてからのことであり、それまでの学術に関する知見の共有が書簡や暗号によっていたことを考えると、学術が本来的にオープンなわけではない¹⁾。とはいえ、学術雑誌を通じ350年以上続いたオープンな伝統なので、オープンサイエンスと聞くとやはり、何を今さら？ という印象を否めない。

経済協力開発機構（OECD）のオープンサイエンスに関わる報告書には、「正式な定義はない」とあり、オープンサイエンスに関連するリソースをまとめているサイト FOSTER には、「研究成果やリソース、研究方法などを共有するための障壁をなくすための動きの総称。学術論文のオープンアクセスやオープンデータ、オープン査読、オープンソース、オープン教材などを

含む」とあり、なおさら、雲を掴むような気分になる。

一方、オープンサイエンスに関わる国際機関や各国の政策文書を吟味すると、「情報通信技術（ICT）」と「新たな学術の手法」が共通のキーワードとして浮かび上がる。例えば、オープンサイエンス政策を強力に推進するEUの政策文書では、「デジタル技術と新しい協働ツールを用いた（中略）、学術プロセスへの新たなアプローチ」とあるし、内閣府の報告書にも「ICTの急速な進展に伴い、オープン化された研究成果やデータを分野、国境を超えて活用することにより、効率的に新たな価値を生み出す」とある。EUが当初、この政策を「Science 2.0」と表現していたことから、オープンサイエンスが、「デジタル時代における新たな学術の次元」としてイメージされていることがわかる。

学術は、学術雑誌という冊子体のかたちを得ることで、知の伝搬性や伝承性を飛躍的に高めた。しかしそれでも、紙媒体であり、持ち運びされる範囲、また流通・出版コストを負担できる主体に、知見の共有は制限された。他方、電子媒体となると、その伝搬性は冊子体以上に広く、さらには情報が瞬時に伝達可能である。しかも流通コストが皆無であるから、研究の解釈の結果としての「論文」だけでなく、研究の素材や生成手法である「データやソースプログラム、関連資料など」も記録として残し、共有することが可能である。

学術は、学術雑誌の誕生により、秘匿性の高い個人研究から、他者の知見に積み重ねることができる研究へと、大きくその性格を変えた。デジタルになり、データも共有できるようになったら、学術はさらに飛躍的な進化を遂げるだろう。オープンサイエンスは、そ

ふなもり・みほ
国立情報学研究所 准教授
〔経歴〕1993年東京大学理学系研究科地球惑星物理学専攻修士課程修了。三菱総合研究所、文部科学省大臣官房国際課、政策研究大学院大学、東京大学（国際連携本部、評価支援室、教育企画室）を経て、2016年から現職。〔専門〕高等教育政策、学術情報流通政策、大学マネジメント。〔趣味〕ヴァイオリン。
E-mail: funamori@nii.ac.jp



うした学術の新たな進化に期待をかけた言葉なのである。

オープンサイエンスの駆動力

オープンサイエンスに、1つの確たる最終型があるわけではないから、その実体を理解するには、いくつかの顕著な動きに着目していく必要がある。

まずICTが要という観点からは、「データ集中科学」に向けた流れがある。曰く、科学は「経験科学」に始まり、「理論科学」、「計算機科学」を経て、現在、第4のパラダイムである「データ集中科学」が出現しつつある²⁾。ビッグデータと聞くと、情報科学分野の話のように錯覚するが、例えば物理的な実験装置からも、データが高サンプリング・レートで機械的に取得され、情報科学以外の分野においても、装置以上にデジタルデータを扱っている時間が、大きな割合を占めるようになった。データ管理や共有、解析が自在な研究環境を提供できる機関は、競争優位に立つ時代である。その観点から、「統合イノベーション戦略」では、「データ基盤の整備」が「知の源泉」として挙げられている。

一方で、こうした理念だけでは新しい考え方は浸透せず、むしろ「社会からの説明責任への要求」の外圧が、オープンサイエンスを推進している。学術研究活動には多額の国費が投じられており、そのアウトプットである論文およびデータについて、オープン化が求められている。研究不正が多発し、研究の透明性が求められるようになったことも、これを後押ししている。

論文のオープンアクセス(OA)化についてはすでに世界的に進みつつあり、2013年のG8で新たに確認されたのは、「公的資金を得た研究データのオープン化の推進」についてであった。公的研究資金を提供するのは研究助成機関だから、これは概ね研究助成条件として課される。論文は公開を前提に執筆されるため問題ないが、データについては、個人情報や知財の観点から取扱いに注意を要する。このためデータについては、公開・共有が「推奨」されるに留まり、代わりに研究データ管理計画(Data Management Plan, DMP)の提出が、現段階では求められている。DMPには、取得

予定のデータの公開/非公開の方針と、データの取得・管理・保存・公開・共有方法などを記述する。

「研究の再現性の危機」に関わる問題も、オープンサイエンスを後押ししている。論文に記載された研究結果に再現性がなかったとする報告事例が相次ぎ³⁾、ネイチャー誌が研究者1500名を対象に2016年に行った調査では⁴⁾、分野にもよるが、研究の再現性がなかったとする研究者が4~8割を占めた。例えば化学では、9割近くの研究者が「他人の研究」を、6割以上が「自分の研究」を再現できなかったとした。研究不正は故意によるものであり、多くの善良な研究者にとっては、これによりデータの管理や公開を求められるのは、迷惑なだけである。一方で、「第三者により再現されるのが、科学である」と学校でも教えられるとおり、「研究の再現性」が担保されないというのは、学術の在り方の根幹に関わる大問題である。

今、データや研究手法、研究環境の伝承を通じた「学術の再構築」が検討されているのである。

オープンサイエンスの実践

オープンサイエンスは、デジタル時代における「新たな学術枠組み」を追求する営みである。これを、研究助成機関から要求される論文のOA義務化やDMP提出の履行で、お茶を濁すのは容易である。しかしそれだけでは、デジタル時代が学術に与える恩恵に浴することができない。ICTにより、論文やデータ、研究環境の共有・伝承可能性が飛躍的に高まった。これまで利用できなかったデータが利用可能となり、これまで存在すらも知らなかった研究者と意見を交わし、世界のどこかで開発された解析ツールも利用でき、また過去の人類の叡智も生で再現可能となるのである。

そのためのデータ基盤は構築されつつある。その可能性を最大限に追求したいものである。

- 1) J.-C. Guédon, "In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing", *Assoc. Univ. Libraries*, 2001.
- 2) T. Hey, "The Fourth Paradigm: Data-intensive Scientific Discovery", *Microsoft Pr.* 2009.
- 3) Nature, "Over half of psychology studies fail reproducibility test", 2015.
- 4) Nature, "1,500 scientists lift the lid on reproducibility", 2016.

© 2018 The Chemical Society of Japan