

# オープンサイエンスをめぐる新しい潮流(その1)

## 科学技術・学術情報共有の枠組みの国際動向と研究のオープンデータ

村山 泰啓 林 和弘

### 概要

データは現代の科学技術・学術にとって重要な存立基盤と言っても過言ではない。研究成果の再検証や、成果の社会との共有、データの再利用による研究活動の発展を図る上で、データの保全・管理と共有・利用は今後、ますます重要なものとなると考えられる。こうした状況下で、科学技術・学術にとっての「データ・マネジメント」が国際的な重要課題となっており、政治の場でも、2013年6月のG8会合で研究活動に係るデータのオープン化が合意事項になった。地球科学を中心にした分野では、世界的な科学データ保存の取り組みが、すでに戦後50年以上にわたり続けられている。近年はさらに、データを原著論文と同等の成果として「出版」して保存し、「引用」するデータ・パブリケーションやデータ・サイテーションの検討、また電子データの保全・管理・利活用に関する図書館の新たな役割の再検討などが課題として議論されている。我が国においても、科学技術・学術研究システムの再構築に、科学技術・学術研究情報の電子化・オープン化を進める制度の整備まで含めて取り組むことが求められる。

**キーワード：**オープンデータ, オープンアクセス, データ共有, ICSU-WDS, CODATA, RDA

## 1 はじめに

2013年のG8首脳会合ではオープンデータ憲章についての合意、あわせて行われたG8科学大臣およびアカデミー会長会合では科学研究データのオープン化に言及した声明が発表された<sup>1)</sup>。

科学技術・学術の研究現場において、近年、論文や発明と同様に重要な成果物、資源として「データ」をとらえるようになってきた。「研究データのオープン化」は、科学技術立国を目指す我が国においても重要な課題である。これは、学術ジャーナルのオープンアクセス化と同じく学術情報共有化の大きな国際世論の流れの中で理解されるべきである。特に日本では、研究論文を中心としたオープンアクセスの議論は比較的盛んであり<sup>2,3)</sup>、2014年3月に日本学術会議のフォーラムも開かれたが<sup>4)</sup>、研究データ

の共有に関する政策的な議論は、科学技術振興機構(JST)において一部見られるものの<sup>5)</sup>、全体的にはほとんど見られない。

オープン化することの情報資源の得失と、オープン化しないことで生ずる科学技術発展上の障害とについては依然慎重な判断が求められるであろう。しかしながら、G8合意にあるように研究データをオープン化することは国際合意である。研究データのオープン化は、研究当事者に閉じず、科学者集団および社会が情報を共有することで、オープンなサイエンスの議論・再検証を可能にしながら、新たな研究機会を増やし、科学と社会の相互信頼を確保しながら研究成果やイノベーションを望めるビジョンと考えられている<sup>6)</sup>。

本レポートでは、近年の国際的な議論の動向を簡単に紹介し、なんらかのオープン化を進めることが妥当な分野・科学技術データないしは学術研究

データセットが特定されたと仮定したときの、社会の科学技術基盤としてのデータのあり方について論ずる。またオープン化することでその分野の科学技術や学術研究がより健全に発展し、より充実した成果をもたらすためにも、背景となる理念を踏まえた上での我が国の対応について基本的な提案を行いたい。

## 2 科学研究の方法論とデータ

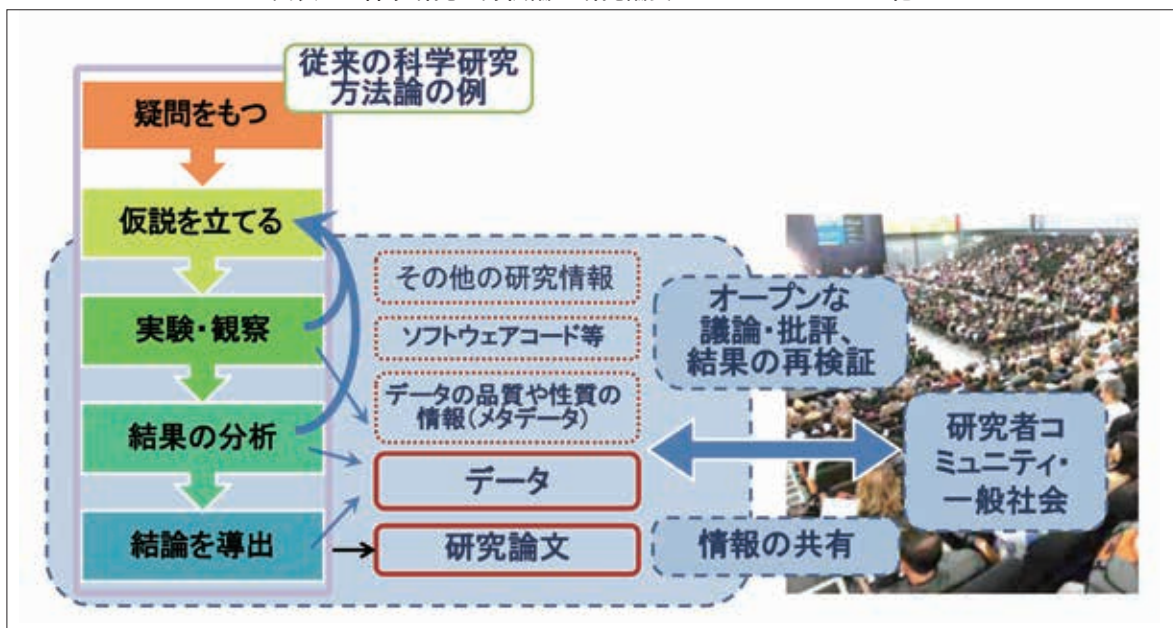
上述の国際的な研究データをめぐる議論では、近代科学の研究方法論が念頭におかれている。科学研究の手続きでは例えば、仮説をたて、証拠に基づく合理的な論理構築により仮説の検証を行い、仮説が支持されれば一定の結論を得る。そこでは結果は例えば専門家コミュニティによる再現性、または統計的有意性の確認により正当と認められる。したがって、第三者による結果の再現性を担保するため、検証可能な記録を共有する必要がある<sup>7)</sup>。論考が基盤となる研究なら、その論理がフォロー可能なように論文に書かれるべきであるし、実験的研究ならば再試験に必要な実験条件が論文に書かれる必要がある。近年は大量のデータの解析結果に基づいて論考した論文も少なくないため、検証にはその論文で使われたデータを入手する必要がある。これにより、同条件での追試験・再検証により研究結果の再現・検証が可能になる。

こうした科学研究成果の再現性に関しては、生命科学分野ではすでに、結論の検証ができない論文が問題視されている<sup>8,9)</sup>。科学が社会と密接にかかわればかかわるほど、社会から科学知、科学的判断の信頼を得ることは重要となり、そこでは科学者集団と社会との情報共有とオープンな検証の担保が重要となるであろうことを示している。そのために、論文流通のオープン化（オープンアクセス）に呼応して、同様に論文のもとになったデータのオープン化が不可欠と考えられている。すべての研究・論文について他者が検証し続けることはできないし、また科学研究の訓練をせずに専門でない人間が十分な成果検証をできるわけではないが、批判や再試験、検証ができる資料の提供機会をつねに維持していることが重要となる（図表1）。さらにオープン化されたデータは検証のためだけではなく、再利用により次の研究を生み出す知的資産として活用される。

## 3 科学技術・学術データの共有をめぐる国際動向

ICSU-WDS（World Data System）は、国際科学会議（ICSU：International Council for Science）の直轄事業・組織として、2008年 ICSU 総会で設立された。また、ICSU-WDS 科学委員会（WDS-SC：Scientific Committee）のもとに、実行組織として国際プログラムオフィスが2012年5月、東京に設置された<sup>10)</sup>。ICSU-WDS の目標は、品質管理された信頼できる科学データを長期間保全することととも

図表1 科学研究の方法論と研究論文・データのオープン化



に、自然科学から人文社会科学まで、多様な分野について相互運用性をもったデータ共有を進めること、等である<sup>11)</sup>。ICSU-WDSの前身となった2つの国際組織、WDC (World Data Centre) および FAGS (Federation of Astronomical and Geophysical Services) はICSUのイニシアチブで50年以上、継続して行われた科学データ保全事業である<sup>12)</sup>。これらの組織は国際地球観測年 (IGY: International Geophysical Year, 1957-1958年) を機に設立されたこともあり、地球科学・宇宙科学が中心であった。

国際組織としてのWDSは現在、データを扱う計算機や大型ストレージシステムを開発・整備することよりも、むしろ世界的な組織間の関係形成 (コミュニティ形成) およびそのコミュニティにおける社会的システムを構築するためのコンセンサス形成が重要なフェーズにある。もちろん国際的なカタログ・メタデータ整備や検索システム開発なども並行しているが、グランドデザインとしては、自然科学から社会科学までの多様なデータを対象とした長期保全と相互利活用を可能とする分散型の世界体制を目標としている。

WDSの作業部会 (WG) の課題の1つにデータ・パブリケーション (出版) がある<sup>13)</sup>。データの出版

は、論文以外にも、科学技術研究開発が社会に還元すべきプロダクトとしてデータを公開するものである。ただし、原著論文出版と比較して考えると、一般読者・ユーザーが信頼できるデータであることを担保すべき査読 (レビュー)、審査基準、印刷するように内容を固定するのか、データ値の修正とその版の管理や引用、など多数の課題がある。WGにおいては科学者・技術者とともに図書館関係者 (ライブラリアン: librarian) や、トムソン・ロイター、エルセビア、ワイリーといったジャーナル出版社がメンバーとなりともに議論を行っている。

他に、ICSU下のより一般的な科学技術データに関係する組織としてCODATA (Committee on Data and Science and Technology: コデータと読まれる) がある。WDSとCODATAは、歴史的経緯もあり、成り立ちや加盟メンバー、投票権、その事業の方向性が異なっている。両者の違いを図表2にまとめている。また、総合的な地球と人類の将来を見据えた学術事業としてFuture Earth事業がICSUや国連などの協同のもとで立ち上がりつつある。WDSおよびCODATAはこれと連携しながら進めるようICSUからも要請されている<sup>14,15)</sup>。

また、G8の下に設けられたGSO (Group of Senior

図表2 ICSU下のデータ関連国際組織 ICSU-WDSとCODATAの違い

	ICSU-WDS (World Data System)	CODATA (Committee on Data and Science and Technology)
<b>設立</b>	2008年 (前身となるWDC、FAGS組織は1957-58年)	1966年
<b>メンバー</b>	データ保有機関、データサービス機関が主体。	各国アカデミー・学術会議、および国際学術連合。
<b>意思決定方法</b>	WDS 科学委員会が最終決定組織。委員は原則的には立候補または各国アカデミー等からの推薦に基づきICSUが決定。	2年毎に開催される総会が最終決定組織。メンバーが投票権をもち、ポリシー策定、決議などを行う。
<b>主な活動内容</b>	品質管理された科学データの長期保全・利用システムの構築、データ公開・出版体制の改善、情報格差の解消など。前身のWDC時代は地球科学分野が大きく貢献してきた。	基礎物理・化学の基礎的な定数・係数の国際推奨値の決定。データ利用・共有に関する国際的なポリシー策定や政策への提言、データ科学の推進。我が国では、物質科学分野などで貢献が大きい。
<b>国際的な貢献</b>	WDC、FAGS時代から50年以上にわたり、科学データ保全事業・情報提供サービスを提供してきた。現在は、自然科学から人文社会科学までのデータの長期保全、利活用体制を検討。	地球観測に関する政府間会合 (GEO: Group on Earth Observations) のデータ共有ポリシーの策定に大きな貢献。これはWDSのデータポリシーへ引き継がれた。

Officials) がグローバル・リサーチ・インフラストラクチャを課題ととらえ、その一部として研究データ・インフラストラクチャ部会が2011年頃から活動してきている。この部会での議論が基礎となつて、研究データに関する新しい国際コンソーシアム RDA (Research Data Alliance) が2013年3月に米・欧・豪の協力によって設立された。RDAへの参加はポリシーに同意すれば原則自由であり、ボトムアップでの会合提案が行われるため、データ共有やデータの相互交換、法的枠組み、海洋データ・遺伝データの相互利用、など多岐にわたる課題についてコミュニティ・コンセンサス形成を目指した議論を行っている<sup>16)</sup>。

なお、欧州委員会内のG8部会担当者によると、オープン・ガバメントと研究データのオープン化とは、別扱いとのことである (Carlos Morais-Pires, 私信 2013)。前者は政府機構の意思決定で情報公開を行うこともできるが、後者の研究データの扱いについては原則、個別の科学者・研究者に委ねられるものであり、科学者や機関のモチベーションを下げず、研究の減速やイノベーションの減退を招かないようにしながら最適なデータ共有へ移行することが、結局は社会のために有用であり重要である。

## 4 研究データのオープン化に対する今後の対応の可能性

### 4-1 研究活動のデータ共有を 発展させるために

• データ・マネジメントのための技術と情報科学・ICT  
科学技術・学術研究データを、原著論文と同等の価値の高い科学技術研究成果として位置づけるとき、そのデータの保存とアクセス性が肝要となる。今日の論文においても約80年前のアインシュタインらによる論文が理論根拠として多数引用されている。これは、論文文献の保存とアクセス機能が長期にわたってよく設計・整備されたシステムであるからこそ可能といえる。一方、論文と同等の高価値な成果物としてとらえられている今日のデータは、80年後といわなくとも、10年後、20年後の我が国の新たなイノベーションを確保するためにも、論文文献の保存・保全と同等のデータ保存の品質、信頼性、期間、アクセス性が担保されなければならない。過去の多数のデータ資源を保存して、80年後の研究開発において必要なときに利用可能にする、そ

うした技術と制度、これをひっくるめて「システム」として機能し続けるようにする必要がある。

現在の電子技術・計算機技術の変革の速さは目を見張るものがある。IP通信や固定ディスクの発明以来、約半世紀ほどの間にさまざまな製品・デバイス・規格が年々改善されて新機種が過去のものを市場から追い出していき、技術と性能が向上してきた。一方、グーテンベルク以来の活版印刷文化に基礎を置く文献の保存、整理、検索、利用のサービスの技術と制度（システム）は、数百年にわたり安定かつ高信頼度で稼働してきた。現在の情報技術、ICT技術がこれまでになかった即時性、高速性、新たなサービス・情報創出などを可能にしながら、活版印刷ベースのシステムと同等の長期的安定性・信頼性を担保した社会の基盤システムとなることで、電子技術文化による10年後20年後、さらには80年後のイノベーションが可能になると期待される。

ところで現状では、情報学・ICT研究と他の科学研究分野とが、非常に密接につながり随所で融合して研究が進展している、とは言い難い。現在では災害対策・防災に不可欠な気象レーダーを例にとると、レーダー工学者と気象学者がそれぞれ協力し合ったというより、両者が歩み寄り、新たな専門性をもつ研究者群が育ち、それぞれの領域から踏み出して、「レーダー気象学」などの新たな研究分野を開拓して初めて成功したとも考えられる。科学技術・学術データの各専門領域と情報学やICT技術はおそらくそれぞれの領域から踏み出した新たな協力関係、新領域をつくることでブレークスルーが実現すると期待される。しかしながら、この新たな連携関係、新領域を日常的に支援できるしくみや予算制度などが乏しいことから、今後は一般の研究者が情報学・ICT分野研究者と連携し、また情報以外の科学分野からの情報学的なアプローチを支援する仕組みが必要となるだろう。

#### • データ保存・管理・利用運用体制

科学技術データ、学術データの保存・管理の議論においては、欧米では特に、図書館関係者（ライブラリアン）が重要な役割を果たしている。国会図書館における資料のデジタル化、電子図書館構想<sup>17)</sup>などのように、電子化時代の図書館の新たな役割が提言され、また、欧米図書館でのデータ収容事業をはじめとして、図書館による文献とならぶ知的情報資産としてのデータの保存・管理事業が新たな図書館の社会的役割として重要視され始めている<sup>18)</sup>。我が国においても、大学図書館や研究機関の図書館は、ジャーナル購読や専門図書閲覧の管理にとどま

らず、文献の電子化やデータ資産の保存という新たな役割と新たな業務の開発を行う、科学技術・学術研究情報の基盤管理・運営組織として新たにデザインされ、我が国の将来のイノベーション創生の機能の一端を担うべき存在にならなければならない。同様に、公的・非営利的な科学技術情報・学術研究情報の保管・管理運用組織においては、図書館と同様に、その情報基盤事業の刷新と新たな基盤を社会に提供するための業務開発が望まれる。

## 4-2 研究成果公開としての研究データの共有を進めるために

### • データ・パブリケーション

論文文献と同様に、データも後年に活用されるべき資料であるならば、原著論文と同様に「出版」されるべき、また査読・編集過程のようにデータの品質評価・出版の過程が必要、との論がある。これには論文査読とは異なったスタイルが必要である。データのフォーマットやメタデータ項目をはじめとする外形的なチェックや、専門家が必要とするデータ品質など、さまざまな議論がされている。求められるデータ品質も、高ければよいというものではなく、利用目的によってさまざまであるし、こうした「データ出版」のスタイルはまだまだ研究コミュニティで検討の途上である。

上記は、データ・パブリケーションという新しいシステム全体の確立へ向けた課題のほんの一部である。コミュニティによるトライ・アンド・エラーを通じたデータ出版活動を、研究費・事業費などの形でサポートしながら支えていく必要がある。

### • データ・サイテーション

データを「出版」して、他の研究でこのデータを使い、また参照して新たな論文を執筆するときに、文献引用・参照と同様にデータもサイテーシ

ョン（参照）することで、同様に、その論文の根拠や関係業績を示して論考を構成する要素となる。また、被引用度数の統計からデータ生成者・提供者の業績をカウントする。論文でいえばサイテーション・インデックスなどに対応する「データ・サイテーション・インデックス」（トムソン・ロイター、2012）などの研究評価での利用が想定される。また論文のPDFと同様にデータにDOI（Digital Object Identifier）を付与して、引用・発見を容易にする取り組みも必要であろう。データを整理して、引用に適した粒度でDOIを付与して、これを論文中で引用する、これに基づき研究成果・データの生産性やデータ生成者の成果を測る、といった一連の活動は、これまでの文献事業でもなければ、データベース構築事業でもない、新たな枠組みであるため、このための活動を資金的・人的・組織的に支援していくことが必要である。

データのパブリケーション・サイテーションを通じた研究評価に関しては、有識者の中には、いずれは例えば研究者の人事審査において原著論文業績リストと、データ業績リスト、およびその被引用度などの評価基準が用いられるような世界になるべきだとの意見もある。巨大科学や大規模実験などで競争力の高いデータを生成することで重要な科学研究成果を出すにあたり、データ生成や整備に寄与した科学者がこれまでかならずしも公平に評価されていないとの声もある。また、このシステムが成立すれば、貴重な科学技術情報を一般公開しても、その生成者・公開者は論文のオープンアクセスと同様に、他の論文からの引用・参照を通じて著者・貢献者として評価されることになる。研究者や機関の利益を損なうことなくデータのオープン化ができるとすれば、そこからデータの相互利用が世界に広がり、新たに公開される膨大な情報資産から、さらなるイノベーションへつながることも期待される。

## 参考文献

- 1) G8 science ministers endorse open access :  
<http://www.timeshighereducation.co.uk/news/g8-science-ministers-endorse-open-access/2004820.article>
- 2) 林和弘. 新しい局面を迎えたオープンアクセスと日本のオープンアクセス義務化に向けて.  
科学技術動向. 2014, 142, p. 25-31 : <http://hdl.handle.net/11035/2475>
- 3) 古西真. 研究基盤としての電子ジャーナル —電子ジャーナルへのアクセスの維持を目指して—.  
科学技術動向. 2011, 119, p.20-27 : <http://hdl.handle.net/11035/2217>
- 4) 学術フォーラム「世界のオープンアクセス政策と日本：研究と学術コミュニケーションへの影響」日本学術会議.

- 2014年3月13日：<http://www.scj.go.jp/ja/event/pdf2/184-s-0313.pdf>
- 5) 日本の科学情報基盤のあり方およびJST情報事業の今後の方向性、科学技術・学術審議会学術分科学術情報委員会（第4回）資料3. 2013年7月24日：  
[http://www.mext.go.jp/b\\_menu/shingi/gijyutu/gijyutu4/031/shiryo/\\_icsFiles/afiedfile/2013/07/29/1338254\\_03.pdf](http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/031/shiryo/_icsFiles/afiedfile/2013/07/29/1338254_03.pdf)
  - 6) Riding the wave – How Europe can gain from the rising tide of scientific data (retrieved August 2014) :  
[http://ec.europa.eu/information\\_society/newsroom/cf/newsletter-item-detail.cfm?item\\_id=6204](http://ec.europa.eu/information_society/newsroom/cf/newsletter-item-detail.cfm?item_id=6204)
  - 7) American Association for the Advancement of Science. (1990) . Science for All Americans. Retrieved from :  
<http://www.project2061.org/publications/sfaa/online/sfaatoc.htm>
  - 8) Boulton, G. (2012) . Why does open data matter and how can we make it a reality? 23rd International CODATA Conference “Open Data and Information for a Changing Planet” . Taipei.
  - 9) Glenn Begley, C., & Ellis, L. M. (2012) . Drug development :  
Raise standards for preclinical cancer research. Nature, Vol.483, p.531, 29 March 2012, doi:10.1038/483531a
  - 10) 情報通信研究機構統合データシステム研究開発室. (2012) . 世界科学データシステム国際プログラムオフィス開所式 :  
<http://isds.nict.go.jp/ceremony01.html>
  - 11) ICSU-World Data System Scientific Committee. (2013) . Constitution of the International Council for Science World Data System (ICSU-WDS) . Retrieved 10 December, 2013, from :  
[http://www.icsu-wds.org/organization/constitution/at\\_download/file1/WDS\\_Constitution\\_06\\_11\\_13.pdf](http://www.icsu-wds.org/organization/constitution/at_download/file1/WDS_Constitution_06_11_13.pdf)
  - 12) 渡邊堯. (2012) . ICSU 世界データシステム (WDS) について . 学術の動向 , Vol.17, No.6, p.6\_11-6\_15 :  
[http://dx.doi.org/10.5363/tits.17.6\\_11](http://dx.doi.org/10.5363/tits.17.6_11)
  - 13) <http://www.icsu-wds.org/community/working-groups/data-publication>
  - 14) Future Earth Transition Team. (2013) . Future Earth Initial Design. Retrieved from :  
<http://www.icsu.org/news-centre/news/top-news/final-report-of-the-future-earth-transition-team-published>
  - 15) 増田耕一, 浦島邦子. 地球環境研究に関する国際プログラムの動向 —Future Earth について—. 科学技術動向 . 2013, 138, p. 18-25 : <http://hdl.handle.net/11035/2427>
  - 16) 恒松直幸, 加藤齊史, 大濱隆司, 村山泰啓. 集会報告 : 研究データ同盟 (Research Data Alliance) 第2回総会 . 情報管理 . P. 724-727, 56 (10) , 2013 : <http://dx.doi.org/10.1241/johokanri.56.724>
  - 17) 国立国会図書館電子図書館中期計画 2004 : [http://www.ndl.go.jp/jp/aboutus/elib\\_plan2004.html](http://www.ndl.go.jp/jp/aboutus/elib_plan2004.html)
  - 18) Karen Antell, Jody Bales Foote, Jaymie Turner and Brian Shults. (2014) Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions. College & Research Libraries, Vol.75, p. 557-574, doi:10.5860/crl.75.4.557

..... **執筆者プロフィール** .....



**村山 泰啓**

科学技術動向研究センター 客員研究官

専門は超高層大気物理学・リモートセンシング。アラスカでの成層圏・中間圏観測に長く携わり、実験観測データベースの開発も行ってきた。ICSU-WDS 科学委員会 ex officio 委員、国立極地研究所南極観測審議委員、京大大学生存圏研究所客員教授、日本地球惑星科学連合・理事などを歴任。



**林 和弘**

科学技術動向研究センター 上席研究官

専門は学術情報流通。1990年代後半より日本化学会英文誌の電子化と事業化に取り組み、オープンアクセスにも対応した。電子ジャーナルから発展する研究者コミュニケーションの将来と、学会、図書館、大学の変革に興味を持つ。