

CA1849 XXXXXXXXXX DataCite : 国立図書館 × DOI × 研究データ

ふくやまじゅり
福山樹里*

1. はじめに

近年、研究データの共有が国際的に拡大しており、DOIを使ってその引用を可能にする仕組みを提供しようとする動きがある。その中心を占める国際データ引用イニシアティブ(DataCite)には、ドイツ国立科学技術図書館(TIB)と英国図書館(BL)が深く関わっている。本稿ではDataCiteと、この2つの国立図書館における研究データに関する取組を紹介する。

2. 研究データの共有に向けた動き

現代において、科学的な研究の成果は、実験や理論構築によって得られた結論を論文の形で公表し、その再現性や統計的有意性が確認されることで正当と認められる。この手続きは、自由な議論と第三者による再現・検証のために必要な情報が共有されていることで担保されている。また、研究成果やその検証に関する情報が社会一般から広くアクセス可能なことで、科学は社会的に信頼性のあるものとして受け入れられている⁽¹⁾。

近年、実験や観察等の過程で生成された数値データ等の研究データは、論文と同等の研究成果として重要視されるようになってきており、これを共有しようとする国際的な動きがある。まず、情報処理技術の進展に伴う国や分野を越えた共同研究の増加を背景に、研究活動の効率化とイノベーションの促進の観点から、研究成果を生み出す材料として研究データを共有しようとする動きがある。また、研究成果の社会への還元という観点から、公的資金の助成を受けた研究は、従来の論文に加え、研究データも成果として公開すべきとする動きがある(CA1818参照)。さらに、研究データの捏造・改ざんや、論文の盗用等の不正行為(研究不正)が社会的に大きく取り上げられることが続く中⁽²⁾、研究の正当性を担保する検証材料として研究データの保存・共有・公開を促す動きも、国際的に共通している。

例えば、G8、経済協力開発機構(OECD)、米国大統領府科学技術政策局(OSTP)等、国際的に影響力のある組織から、公的資金による研究成果は研究データを含めて原則として公開するよう指示が出されている。日本では、内閣府の「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」の報告書「我が国におけるオープンサイエンス推進のあり方について」

て」⁽³⁾や、日本学術会議の研究データの保存の必要性に言及した文書⁽⁴⁾⁽⁵⁾の公開等の動きがある。

3. DataCite 創設までの 10 年

研究データの共有の動きと並行して、科学技術データ委員会(CODATA)、DataCite、研究データ同盟(RDA)(E1531参照)等、研究データの活用を推進しようとする国際的な組織による取組がある。中でも、DataCite⁽⁶⁾はその創設に2か国の国立図書館が関与している。DataCiteは、デジタルオブジェクト識別子(Digital Object Identifier: DOI)(CA1836参照)を研究データに付与する登録機関である。DOIは、これまで電子ジャーナルをはじめとする論文等に対して、リンク維持のために付与されてきた。DataCiteは、DOIを使って研究データの引用と適切な活用を可能にすることを目的としている(E1537参照)。以下、DataCiteの創設の経緯を、TIBのブラーゼ(Jan Brase)氏らによる記事⁽⁷⁾から紹介する。

DataCiteは2009年12月に設立されたが、その始まりは、2000年頃にCODATAのワーキンググループ(WG)「一次データの引用可能性(Possibility of Citing Scientific Primary Data)」が中心となって行った議論と分析にある。通常、研究の過程で生成されたデータ(一次データ)は、その研究者や研究機関によって管理されている。一次データを活用するためには、そのデータの生成条件を記述したドキュメントが必要であり、長期にわたって管理されていなければならない。しかし、研究成果の検証や別の研究目的での再利用等を考慮して研究データを管理することは、研究者にとって手間のかかる仕事である。また、研究データは論文のように引用する習慣がなく、研究成果として敬意を払ったり業績評価の対象としたりする習慣もなかった。そのため、適切な管理が行われていなかった。

このような分析の結果、2002年のWGの最終報告書では、一次データを論文に引用することでデータの入手可能性を向上させるという考えが示された。2003年の論考⁽⁸⁾では、研究データを引用可能にする方法として、永続的な識別子を研究データに付与することと、論文の流通において普及しているDOIを識別子として使うことを推奨している。また、その実施主体としてTIBとともに、世界データセンター(World Data Center: WDC、現ICSU-WDS)等のデータアーカイブ機関を挙げている。

これらを踏まえ、TIBとWDC等は2003年から実装を試行、2004年にTIBが世界で初めて研究データにDOIを付与し、2006年からはサービスを提供するためのシステム開発と改修を行った。その後、2007年にプロジェクトの報告をすると、2008年にはBL

*電子情報部 電子情報流通課

を含む欧州の図書館等が関心を寄せ、2009年には米国、カナダ、オーストラリアにも波及し、国際的な組織としてDataCiteが創設された。2015年3月時点の会員数は、TIB、BL、オーストラリア国立データサービス(ANDS)、ジャパンリンクセンター(JaLC)等、22機関である。なお、国立国会図書館(NDL)はJaLCの4つの共同運営機関の一つである。JaLCでは、現在、研究データへのDOI付与に関する課題等を検討する実験プロジェクトを実施している⁽⁹⁾。

4. TIB と研究データ

ドイツには、科学技術情報、医学・農学、経済学の各分野を専門とする3つの国立図書館がある。科学技術情報を扱うTIBは、ドイツ国内のみならず国際的にも、研究コミュニティと産業界に対してあらゆる科学技術情報へのアクセスを提供することを使命としている。そのため、情報処理技術の進展したデジタル時代においては、提供する情報の範囲も従来の論文や書籍に留まらず、一次データ、動画、ソースコード、シミュレーションデータ、スライド資料等、あらゆる種類の情報に拡大する必要がある。また、TIBでは、利用者の求める情報自体を保有していない場合でも、その所在情報を把握・提供することが未来の図書館の仕事になると考えている。これを実現するには、世界中の科学技術分野の研究データに関する情報を登録し、当該コンテンツの記述を取り出せる形で、図書館の目録情報内に持つ必要がある⁽¹⁰⁾。

TIBはDOIの登録機関として、世界中の非営利の情報機関や図書館が、より容易にDOIにアクセスできるようにすることを目指している。このため、TIBは、科学技術研究の国際性と、生成された研究データが各研究機関や研究者によって個別に管理されている現状を踏まえ、ドイツ国内の他の2つの国立図書館を含む、国内外の国レベルの代表的な機関と、分野横断的にグローバルな協力体制を構築しようとしている。その際、データの保存・維持管理、データの品質の保証・評価、メタデータの作成については研究機関やデータセンター等のデータ提供者が責任を持ち、TIBは、そのメタデータを蓄積・保存して検索可能にする役割を担うこととした⁽¹¹⁾。

現在TIBは、ウェブベースのポータルであるGetInfo⁽¹¹⁾を通じて、一次データ、化学構造等の化学物質情報、建築の設計モデル、論文の全文データ、目録データ等、多様なデータへのアクセスを提供している。非テキスト情報にとって重要なことは、目録と外部コンテンツの安定したリンク関係である。GetInfoでは、研究データの多くにDataCiteのDOIが登録されているため、検索結果からDOIを経由してデータ

提供元へアクセスすることや、データを単体で引用することができる。また、そのデータを活用した論文があれば、論文へのリンクをたどって研究成果を検証することもでき、同じ課題の多様な側面についての共同研究を促進することにもなる⁽¹³⁾。TIBはDataCiteを主導するだけでなく、研究データの管理に必要なインフラの提供を目的として、カールスルーエ工科大学やライプニッツ協会等とともに“RADAR”⁽¹⁴⁾というリポジトリを運営している。ここでは、研究データの保存・公開、論文とのリンクの増加を目指している。このように、研究データの管理と活用のためのインフラ構築や、非営利組織の国際的な連携体制の構築を主導することで、TIBは、あらゆる科学技術情報へのアクセスを提供するという使命を果たそうとしている⁽¹⁵⁾。

5. BL と研究データ

BLは、2008年以降、研究データを「国の財産(National Asset)」と位置付け、保存とアクセスを保証すべき対象を、従来の物理的な媒体から研究データにまで拡大した⁽¹⁶⁾。2008年から2011年にかけての戦略(The British Library's Strategy 2008-2011)⁽¹⁷⁾では、「5. デジタル情報のインフラ構築」の一環として、他機関と協力の上、研究データセットプログラム⁽¹⁸⁾を含む研究データ関連事業を推進することとし、続く2011年から2015年の戦略(Growing Knowledge: The British Library's Strategy 2011-2015)⁽¹⁹⁾では、「2. 研究をしたい全ての人へのアクセスの確保」として、全分野のデータセットへのアクセスや引用・再利用の促進等を挙げている。さらに、最新のビジョン(Living Knowledge: The British Library 2015-2023)⁽²⁰⁾では、「2. 研究」として、大規模なデータ分析におけるイノベーションを推進し、効果的な研究支援のためにBL自身の研究力を向上する等としている。

BLでは、研究データセットを、組織化されたデータコレクションと定義し、図書館の内外にあるデジタル形式のものに焦点を絞った。この定義では、画像、音声、数値データが含まれるため、デジタル化した図書、新聞、地図、目録データ等、データセットに当たるものをBLは既に扱っていたと言える。また、国内には、英国データアーカイブ(UK Data Archive)のようにデータの保存を担う機関も存在する。しかし、検討の結果、ネットワーク上に存在する研究データの保存や永続的な同定・識別は、BLが取り組むべき新しい課題だと位置付けた⁽²¹⁾。

この課題解決のため、BLは2008年からDataCiteの前身のプロジェクトに参加し、2009年のDataCiteの創設時に会員となった。2010年には、国内のデータ管理機関とともにDOI登録プロジェクトを試行し、

2012年から11機関を対象にDOI登録サービスの本格的な提供を開始した。

現在、英国内の各分野の国レベルのデータ管理機関を対象に、DOI登録のインフラを提供すると同時にワークショップを継続的に開催している。その目的は、関係機関との協力関係の維持とコミュニティの構築である。研究データへの永続的なアクセスの保証には、各機関におけるDOIの適切な維持管理が必須である。そのため、ワークショップでは、各機関がDOIの永続性に責任を負うことを説明している。また、DOIの登録手順のデモンストレーションといった基本的な情報を提供するだけでなく、参加機関における事例報告も行っている。報告には、例えば、DataCiteへのメタデータ登録によって、随時更新される研究データのバージョン情報が管理でき、他の研究での再利用が可能になることや、当該データを一定期間非公開にする必要がある場合でも、DOIの付与とメタデータの登録によってデータの引用が可能になるため、研究の信頼性・透明性の確保につながること等、DataCiteに参加する利点に関する情報も含まれている。こうした情報を、まだDataCiteに参加していない機関も含めて共有し、意見交換を行っている⁽²²⁾⁽²³⁾。このような普及・啓発活動とコミュニティ形成によって、研究者によるデータセンターへのデータ提供とデータの引用を促進することで、BLはその使命を果たそうとしている。

6. おわりに

研究データと論文等の関連情報が相互にリンクされ、容易に引用できれば、研究活動の効率化や研究成果を検証することが可能な環境を提供することにつながる。しかし、多くの研究が時限的なプロジェクトとして実施され、多くの研究者がより良い研究環境を求めてしばしば所属を変える現在においては、研究データの永続的な管理とアクセスの確保を研究者のみの仕事とすることは、現実的ではないだろう。また、有用な研究データへの永続的なアクセスを保証するためには、永続性を担保する意思と体制を持つ機関が関与し、識別子とそれに対応する情報とのリンクの維持管理を適切に行う必要がある。さらに、研究活動の国際性を考慮すると、国や分野を越えた国際的な連携も欠かせない。これらを考え合わせると、DataCiteを通じたTIBとBLの活動は、国内の情報ニーズに応える機関である国立図書館が担うべき重要な役割を果たしている一例だと言えるだろう。NDLでは、「私たちの使命・目標 2012-2016」⁽²⁴⁾において、知識・文化の基盤となることを使命とし、「目標3：情報アクセス」で、必要な情報に迅速かつ的確にアクセスできるように、新

しい情報環境に対応して、利用環境を整備することを挙げている。このことから、NDLもまた、研究データも含めて、情報提供の基盤となる識別子の永続的な維持管理に取り組んでいく必要があるのではないだろうか。

- (1) 村山泰啓, 林和弘. 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. 科学技術動向. 2014, (146), p. 12-17. <http://hdl.handle.net/11035/2972>, (参照 2015-03-30).
- (2) 小林信一. 我々は研究不正を適切に扱っているのだろうか (上) 研究不正規律の反省的検証. レファレンス. 2014, (764), p. 25-45. doi: 10.11501/8752135.
- (3) “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”. 内閣府. <http://www8.cao.go.jp/cstp/sonota/openscience/>, (参照 2015-03-30).
- (4) 日本学術会議. 報告 オープンデータに関する権利と義務: 本格的なデータジャーナルに向けて. 2014, 17p. <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-h140930-3.pdf>, (参照 2015-03-30).
- (5) 日本学術会議. 回答 科学研究における健全性の向上について. 2015, 29p. <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-k150306.pdf>, (参照 2015-03-30).
- (6) DataCite. <http://www.datacite.org/>, (accessed 2015-03-30).
- (7) Jan Brase, Irina Sens and Michael Lautenschlager. The Tenth Anniversary of Assigning DOI Names to Scientific Data and a Five Year History of DataCite. D-Lib Magazine. 2015, 21 (1/2). doi: 10.1045/january2015-brase.
- (8) Michael Lautenschlager and Irina Sens. Konzept zur Zitierfähigkeit Wissenschaftlicher Primärdaten, 2003-08-19, 11p. http://mms.dkrz.de/pdf/klimadaten/service_support/Publications/Paper-Konzept-Primdaten.pdf, (accessed 2015-03-30).
- (9) ジャパンリンクセンター. <http://japanlinkcenter.org/top/>, (参照 2015-03-30).
- (10) Jan Brase. DataCite and linked data. JLIS.it. 2013, 4 (1), p. 365-373. doi: 10.4403/jlis.it-5493.
- (11) Jan Brase. Research Data in Library Catalogs. Defense Technical Information Center. <http://dtic.mil/dtic/annualconf/ResearchDatainLibraryCatalogs.ppt>, (accessed 2015-03-30).
- (12) GetInfo. <https://getinfo.de/>, (accessed 2015-03-30).
- (13) Brase 2013. op. cit.
- (14) RADAR. <http://www.radar-projekt.org>, (accessed 2015-03-30).
- (15) Uwe Rosemann. Textual and non-textual objects: Seamless access for scientists. http://haxel.com/icic/2013/Programme/monday-14-oct-2013/text-and-non-textual-objects-seamless-access-for-scientists/at_download/attachfile, (accessed 2015-03-30).
- (16) J. Max Wilkinson et al. British Library Dataset Programme: Supporting Research in the Library of the 21st Century. LIBER Quarterly. 2010, 20 (1), p. 94-104. <http://liber.library.uu.nl/index.php/lq/article/view/7979/8288>, (accessed 2015-03-30).
- (17) “The British Library’s Strategy 2008 - 2011”. British Library. <http://www.bl.uk/aboutus/stratpolprog/strategy0811/strategy2008-2011.pdf>, (accessed 2015-03-30).
- (18) “Datasets Programme”. British Library. <http://www.bl.uk/datasets>, (accessed 2015-03-30).
- (19) “Growing Knowledge: The British Library’s Strategy 2011 - 2015”. British Library. <http://www.bl.uk/aboutus/stratpolprog/strategy1115/strategy1115.pdf>, (accessed 2015-03-30).
- (20) “Living Knowledge: The British Library 2015 - 2023”. British Library. <http://www.bl.uk/aboutus/foi/pubsch/pubscheme3/living-knowledge-2015-2023.pdf>, (accessed 2015-03-30).

- (21) Wilkinson, et al. op. cit.
 (22) Elizabeth Newbold. Working in Collaboration with Data Centres. DataCite. http://datacite.inist.fr/IMG/pptx/ewbold_august25.pptx, (accessed 2015-03-30).
 (23) Working with the British Library and DataCite: Institutional Case Studies. http://www.bl.uk/aboutus/stratpolprog/digi/datasets/DataCiteCaseStudies_2013.pdf, (accessed 2015-03-30).
 (24) “「私たちの使命・目標 2012-2016」及び「戦略的目標」”. 国立国会図書館. <http://ndl.go.jp/jp/aboutus/mission2012.html>, (参照 2015-03-30).

[受理：2015-05-11]

Fukuyama Julie

DataCite: National Libraries×DOI×Research Data

CA1850

『日本十進分類法』新訂 10 版の概要

たかはしりょうへい*
高橋良平*

1. はじめに

『日本十進分類法』(NDC)は、標準分類法として日本の図書館に広く普及している図書分類法である⁽¹⁾。1929年の刊行以来版を重ね、この度、新訂10版(以下、10版)。他の版についても同様)が2014年12月に刊行された。9版刊行から実に19年ぶりの改訂である。筆者は、国立国会図書館(NDL)収集書誌部国内資料課で国内刊行図書の目録作成に従事する傍ら、2010年から日本図書館協会(JLA)分類委員会委員として10版改訂作業に携わってきた。本稿では10版の改訂内容について概観し、併せて分類委員会の今後の予定について触れる⁽²⁾。なお、本文中の意見にわたる部分は、分類作業実務者である筆者の個人的見解であり、所属機関を代表するものではないことをお断りしておく。

2. 改訂方針

2004年に公表された改訂方針⁽³⁾では、9版の改訂方針を継承して書誌分類を目指すものの、「実務の継続性の確保の観点から」、NDCの根幹に関わる体系の変更は行わないことが確認された。そのため分類委員会では、既存の分類体系の枠内で、改訂方針が目指す「分類作業が行いやすく、また利用者にも分かりやすい分類表」を実現すべく検討を重ねてきた。

3. 主な改訂点

3.1 「序説」・「使用法」

10版では、9版の「解説」を廃止し、NDCの基本構造を説明する「序説」と、分類作業の進め方について解説した「使用法」を新たに設けた。また、「各類概説」を独立させ、「序説」及び「使用法」の説明に沿って各類の構造が理解できるように内容を拡充した。文中に登場する専門用語は「用語解説」にまとめている。

今までNDCに関する概説書やテキストは多数出版されてきたが、これらと同じ水準の説明がNDCの本体に組み込まれ、NDCの仕組みや考え方を分類委員会として公式にまとめた意義は大きいといえる。

(1) 序説

多様な主題をNDCの分類体系に位置づけるには、資料の「主題」と、主題の扱われ方である「観点」を把握する必要がある。しかし、その重要性にもかかわ

*収集書誌部国内資料課