

## CA1875

研究データ共有の国際動向と  
図書館の役割：RDA 第7回総会報告むらやまやすひろ\*  
村山泰啓\*

2016年3月1日から3日までの3日間、RDA (Research Data Alliance: 研究データ同盟) の第7回総会 (7th Plenary Meeting) が、一橋講堂 (東京都千代田区) において行われた。またこの前日、2月29日には科学技術振興機構 (JST) 主催のデータシェアリングシンポジウムが同じ場所で開催された<sup>(1)</sup>。

各学問分野の研究活動を通じて生み出される研究データを重要な情報資産ととらえ、その情報整備、データ共有を通じてオープンサイエンスなどの活動を推進しようという動きが国際的に広がり、我が国の第5期科学技術基本計画<sup>(2)</sup>においてもオープンサイエンスの推進が書き込まれている。

ここでいうサイエンスは、自然科学・技術分野のものだけではなく、社会科学・人文科学分野までふくむものである。RDAがScientific DataでなくResearch Dataとする理由については、以前、RDAの組織命名の経緯を聞いた際、企業活動も含めて調査、研究に関わる知的価値の高いデータをすべて対象にしたいからと聞いた記憶がある。

学術情報流通の視点からも、今後は文献にとどまらず幅広く多様な情報資産を、研究コミュニティおよび国内・国際社会で共有・流通し、有効活用することが極めて重要だと考えられている<sup>(3)</sup>。

しかし研究データを生み出す現場では、管理ノウハウが不十分なことも少なくないし、書誌のような情報流通のための情報整備も研究者は不得意なことが多い。図書館、博物館などの専門家による情報の管理・組織化の手法、技術、経験、スキルは今後非常に重要になると思われる。RDAにおいて創設当初から図書館員の登録・参加が少なくないことは、国際的にこの分野で図書館・図書館員に期待される役割、図書館員自身の描く将来像が重要であることを示している。

本稿では、これらの会合の参加報告とともに、RDAという国際組織の現状と今後の展望について概観してみたい。

## 1. RDAとは

RDAは2013年、G8科学技術大臣会合の研究データオープン化の声明に前後して立ち上げられた、研究データの利活用を目指す国際組織である<sup>(4)</sup>。欧州と米

国での議論がその発端と言われるが<sup>(5)</sup>、グローバルな活動を本格化させたきっかけが、2013年の国際政治の場での議論であったことは無視できないだろう。日本に対して英国や全米科学財団 (NSF) 等からRDAへの参加・協力要請があったこともあり、我が国では国際動向を踏まえたオープンサイエンスへの対応、という形で議論が進められてきた<sup>(6)</sup>。

事務局運営の資金は、米国、欧州、豪州が少額ながら支出しているとのことである<sup>(7)</sup>。またRDAと並行して当初、G8国および関係6か国 (O6) の政府関係者・研究資金配分機関などが参加するG8+O6データ・インフラストラクチャ・ワーキンググループにおいてRDAの将来や運営に関する議論も行われてきた。これはRDAコロキウムという組織を経て、CRE (Collaborative Research e-Infrastructure) という新たな体制にすべく検討・調整が進められている。

RDAの総会は2013年に第1回が開催され、以来、年2回の総会開催にはIT (情報技術) / ICT (情報通信技術) 技術者、情報学者、個別分野の科学者 (例えば遺伝学、海洋学、地球惑星科学、農学、天文学等)、図書館員、データセンター/リポジトリ管理者、データキュレーター、研究資金配分・提供機関、政府関係者など多様なコミュニティの人材が参加している。

過去の総会参加報告や研究データ共有に関する考察 (CA1818, E1566, E1676 参照)<sup>(8)</sup>にあるように、RDAは権限・強制力ではなく参加者からのボトムアップの提案にもとづき、コミュニティ・コンセンサスをベースにして研究データ共有のポリシー、方式・制度、技術、基盤システムなどについて検討、合意形成、実践していくことを目指している<sup>(9)</sup>。RDAへの個人加入は、ウェブサイトから自由にできるようになっており、SNSのように容易な登録が可能である (事務局によると2016年5月16日現在、個人登録会員数は3,976名とのことである)。

しかしRDA総会に初めて参加する人にとっては、何を議論しているのか大変わりにくいことが多いようである。RDA総会ではシンポジウムで序章・背景説明から語るような講演は少なく、部会で具体的な問題提起やその議論を行うことが主である。つまり数多くの委員会・検討会合の集合体ととらえたほうがわかりやすい。部会の日ごろの活動・議論を総会であらためて整理してまとめ直すものと考えれば、継続的に参加しないと背景・文脈が捉えにくいのもうなずける。しかし理解できる部会が見つかり、国際的に検討中の課題について、進行中の議論を目の前で見て参加することができる。いわば国際社会において「熱い鉄を叩いて整形している」現場である。それは興味深いし、ときにはエキサイティングである。

\*情報通信研究機構統合ビッグデータ研究センター

## 2. データシェアリングシンポジウム報告

このシンポジウムは、サブタイトルを「科学の発展への起爆剤～データ駆動型科学の推進に向けて～」として開催された。午前中は主にポリシー・政策推進面からの展望を中心にスピーチ、講演があり、オープンサイエンスや研究データ共有の重要性と、それらを推進する立場を確認した。午後は情報学、地球科学、物質科学、バイオサイエンス、脳科学、人工知能の各分野の専門家からの報告やパネル・ディスカッションが行われた。登壇者から「そもそもオープンサイエンスのメリットは？」といった本質に立ち戻る疑問もあわせて提示されるなど重要な論点について議論された。

## 3. RDA 第7回総会：全体像

RDA 第7回総会（以下 RDA-P7）の総参加者数は373名であった。過去の総会参加者は欧州内で500～600名程度、米国内で400名弱であったので、過去の総会と同規模といえる。これまでRDA総会は欧州か北米で行われ、それ以外の地域からの参加者は少数派であり、議論を行う主体がどうしても欧州・北米・豪州が中心になる傾向があった。今回、欧州・北米外ではじめての開催地が東京となり、日本から100名を超える参加者があったことは意義深い。

参加者の国・地域別内訳は、日本113名、欧州128名、米国71名、豪州11名等であった。これまでアジアからは、日本をのぞけば参加があっても1国1～8名であった。この傾向は今回も変わらなかった<sup>(10)</sup>。ただこれまで参加のなかったマレーシア、フィリピン、シンガポールから1～2名ずつ参加者があった。

RDA総会初日の全体セッション（Plenary session）では、しばしばホスト国の責任者や代表的な人物、閣僚・アカデミー代表等からスピーチがある。RDA-P7では内閣府総合科学技術・イノベーション会議議員の原山優子氏、文部科学省大臣官房審議官の岸本康夫氏、科学技術振興機構上席フェローの大竹暁氏らがいさつを行い、欧州委員会（EC）からは官房長官のMichael Hager氏のビデオメッセージが流された。

Hager氏の話を受けて、その後のパネルセッションはECで現在検討中の欧州オープン科学クラウド（European Open Science Cloud: EOSC）<sup>(11)</sup>施策提案を主眼にした議論が行われた。EC閣僚へ政策提言を行う高級専門検討会（High Level Expert Group）座長B. Mons氏のほか、情報通信技術分野からS.Wolff氏（Internet2主幹科学者）、データ利活用分野からはM.Parsons氏（RDA国際事務局長）ら主要人物が参加したことは興味深い。欧米では研究データ基盤<sup>(12)</sup>整備事業はすでにいくつか立ち上がっている<sup>(13)</sup>。EOSCでは世界全体の研究データ基盤、リポジトリ等の連携

が不可欠との意見が出てきており、筆者も登壇して我が国の研究データ基盤および関連する事例を紹介した。2日目の全体セッションでは午前中、RDAの提言・アウトプットとして7つのテーマが語られた。午後、喜連川優氏（国立情報学研究所長）の基調講演では、研究データ基盤技術開発や研究データ利活用事例などが紹介された。

## 4. RDA 第7回総会：部会活動

RDAの特徴は上述したようにボトムアップで提案される部会にある。部会は主にWG（Working Group）、IG（Interest Group）、BoF（Bird of Feather）に分類される。今回開催された会合は、WG会合8、IG会合25、BoF会合10、合同部会会合9であった。

RDA活動の中心となるWGからは、近年の総会で毎回、複数の報告、提言が出されており、RDAで設計された課題解決の「エコサイクル」が稼働している様子がうかがえる。

「エコサイクル」とは以下のようなものである。BoFあるいは非公式グループによる部会立上げ、基本テーマ討論から、IGによる共通認識形成、課題の洗い出し、解決へのアプローチなどの検討をへてWG提案、WG設置後は18か月のアクション（会議開催や実験・実装など）をへて報告・提言を行う。

わずか18か月で解決できる課題は限られる。例えばデータ引用、データリポジトリ整備、データパブリッシングなどをはじめとして、1～2年の検討や実験では解決できない規模の課題も多い。各WGは18か月で解決できる課題にテーマを細分化・ブレイクダウンするなどして、結果を出していく。一方、例えばデータパブリッシングIGのように複数のデータパブリッシング関連WGの連携拠点のように見えるIGもある（WGには存続期限があるがIGはない）。オープンサイエンスにおけるデータ取扱いに係る課題の多さ・複雑さを考えれば、18か月ごとに個別課題を着実に解決しながら、段階的に大規模な課題の解決へアプローチしていくようなグループもあってよいわけである。

部会の課題は、ルール、作法、実践方法、研究者へのインセンティブ、技術基盤などに及び、それぞれ分野総括的な課題と個別分野ごとの課題がある。課題によっては、法令など多くの側面について部会間で協調・調整しながら検討を行っている。

## 5. 外部機関との協働、コミュニティ形成

RDAへの参加者のうち、部会やRDAの運営に積極的に関与する人々は、自身の出身国や、他の国際組織・プロジェクトでも活躍する有識者が多

いように見受けられる。筆者の知る範囲でも ICSU-World Data System (WDS)、Committee on Data for Science and Technology (CODATA)、Data Seal of Approval (DSA)<sup>(14)</sup>、DataCite (CA1849参照)、ORCID<sup>(15)</sup> (CA1740 参照) 等、図書館、博物館、計算機工学、個々の科学領域のコミュニティは RDA 活動の本質的な部分に重要な情報と人材を供給している。見方を変えれば、RDA の本質が、さまざまなステークホルダーが知恵を寄せ合い新解決法を探るプラットフォームであるとすれば、人材や機関間・個人間ネットワークが共通しているのはごく自然なことである。

例えば WDS や DataCite などはデータ保全・管理、データへの DOI 付与、といった個々の課題を中心に活動をすすめる「プレイヤー」であり、個別の知識や実践例を作り出すコミュニティであるが、RDA はそうした複数のコミュニティ同士が結びつくための舞台ともなっている。RDA はそういう意味での国際プラットフォームとして機能しはじめていえるだろう。

RDA 総会は、RDA によって公認されたサイドイベント（今回、公式には co-located event と呼ばれた）とともに開催されることが多い。今回、サイドイベントの開催に係わった国際組織には、CODATA、OECD、EUDAT (European Data Infrastructure)<sup>(16)</sup>、RDA/WDS データパブリッシングのグループ、EDISON<sup>(17)</sup> などがあった<sup>(18)</sup>。また RDA-P7 の直前に OECD Global Science Forum (GSF)<sup>(19)</sup> は CODATA と ICSU-WDS という 2 つの国際データ組織と協同で 2 つのプロジェクトを立ち上げた。RDA-P7 が、その最初の顔合わせの場選ばれたわけである。

また日本がホストする RDA-P7 において新たな機関リポジトリに関するセッションやサイドイベントが開催され、そこで日本人の情報学研究者が座長など重要な役割を担って参加していた。これまで日本からのこの分野の関与があまりなかったことを考えれば重要なステップと言えらるだろう。

## 6. 情報基盤システムにおける紙メディアから電子メディアへの転換

研究論文・文献の多数が PDF ファイルなどの電子ファイルとして流通する現在、さらに研究データ、政府機関などの公的データをはじめとして、インターネット上の多様な情報資産（ここでは、オンラインで取得またはアクセス可能なデジタル・オブジェクト全般を情報資産とする）の有効活用が望まれている<sup>(20)</sup>。蓄えられた静的なデータだけでなく、動的なデータストリームまで考慮すれば、IoT (Internet Of Things) など先端技術はさらにサイズが大きく豊かな情報資産を生み出す可能性がある。

一方、研究活動の結果として、データは論説・知識としての学術文献などの新たな価値の高い情報資産を生み出す。その利活用は科学・技術における知的・社会的・経済的發展にとって欠かせないものの一つだったとあってよい。これまでの成功は、主に印刷・紙メディアの技術と文化を基盤として成し遂げられてきた。これまでの科学研究および学術情報流通にとっては、国際的な図書館間の連携システムや文献の編集・出版システム、そこでのキュレーション・査読といった洗練された仕組みが不可欠な基盤であった<sup>(21)</sup>。これは、技術システムと社会（制度的）システムが相互に調和した総合システムとして機能していると考えられる。このような総合的なシステムとなっはじめて、社会で広く活用される基盤となったと言ってもよいであろう。

現在進められている、Science2.0 やオープンサイエンスといった議論では、印刷技術・紙メディアに代わって電子情報通信技術・電子メディアが重要な役割を果たす。そこでは、究極的には、上述の印刷・紙メディアで構築されてきたような洗練された総合システムが、電子情報について実現する必要がある。

RDA は、国際的に進められている、こうした電子メディア上の技術システム、社会システム、および両者の結合系としての総合システムを、構築・確立しようとする営為の一部であると解釈することもできる。RDA のいくつかの部会活動は、このような壮大なグローバルシステムの構成要素をつくらうとしているように受け取れる。

言い方を変えれば、RDA は IT/ICT の工学研究・技術開発から、広い意味での業務実施者・利用者層全般（印刷文化における著者層・編集者層・査読者層・出版者層・購読者層・再利用者層、など多層構造を構成する）を巻き込んで、前段で述べたような技術システムと社会システムをつくりながら、総合システムの形成プロセスとなろうとしている。現在は全貌があまり明確でなくても、システムの最適化へ向けた壮大な基本設計作業、またそれに必要な技術要件・利用要件の洗い出し、必要な構成要素（技術的、社会的、制度的）をかたち作っていく活動ともいえるだろう。

図書館は、これまでは紙メディアを中心に情報資産の管理・保全・サービスを提供し、図書館員はこれを支える専門家として、社会から信頼を勝ち得てきた。科学技術や学術に関する活動の基盤となるような情報流通システム（情報流通基盤システム）の歴史の転換点であるかもしれない現在、今後はどういった情報資産、どのような業務を対象にすべきか、紙メディアおよび電子メディアの共存共栄のあり方、そこで求められる図書館、および専門家としての図書館員の役割は

何なのかについては、今後も辛抱強く考え検討される必要があるだろう (CA1825, CA1851 参照)。もしも、大量かつ多様な紙および電子メディアを扱う広い意味での電子図書館システムが実現して、図書館が電子的な情報流通基盤システムの一部となり、社会に豊かな情報資産を提供できるようになるならば、それはオープンサイエンスやオープンイノベーションによる新しい豊かさを享受できるようになった社会にとって、これまで以上に不可欠な存在となっているのではないだろうか。

## 7. おわりに

本稿では、RDA-P7 の開催報告に関連して、RDA に関わる研究データの問題について、図書館の関わりを念頭において述べた。研究データ共有やオープンサイエンスは大きな概念であり、例えば、科学・学術の進展、情報基盤システムの技術研究・開発、科学と社会の制度改革、社会経済の発展、などさまざまな切り口で語ることができる。その中で本稿では、電子的に取り扱われる研究データ、および、上述のような学術情報流通基盤に用いられる技術の変化の可能性、そこでの図書館への期待、に主な焦点をあてている。今後、電子データについて情報流通基盤システムを活用する体制が実現していく過程において、図書館にも重要な機能を期待されることは国際的にも、おそらく間違いないだろう。

我が国において RDA やオープンサイエンスなど、電子的な情報資産の取り扱い、組織化などの議論がどのような成果をもたらし、新たな情報流通基盤システムがどのように実現していくのかはまだわからない。RDA の内部エコサイクルをはじめさまざまな取り組みの中で生まれていくデータマネジメント技術、方式、制度に、日本がどのように関わっていけるのか、今後も注視していきたい。欧米の中には国策として、電子データに関するオープンサイエンス推進や電子情報流通基盤整備を開始しているところもある。これは科学技術分野の対応だけにとどまらず、例えば EC の欧州デジタル単一市場戦略 (Digital Single Market Strategy)<sup>(22)</sup> ではデジタル経済戦略によって社会、産業界を変革していこうという方向性を示し、その中で、必要なオープンサイエンスの推進方策を位置づけている。

社会が電子情報を利用する体制へ移行していく流れは避けられないのではないかと筆者は考えている。この潮流の中で、国際的に研究データの取り扱い方、ルール作成・標準化が完了してしまうと、後から日本の意見を反映するのは極めてむずかしいであろう。国際的にもまだ方法論、実践方法が議論されている現在、今

ならまだ日本も国際的議論に参加できる、追いつけるかもしれないという意見もある。図書館、博物館、大学・研究機関、企業、予算配分機関、行政機関、といった多様なステークホルダーがともに国際動向の変化をフォローしながら、また相互に意見交換しながら、我が国にとって有益な将来像や独自の知見を作ってゆく、そういう輪が生まれることが期待されている。

なお RDA-P7 ののち、RDA 参加報告と議論を行う「研究データとオープンサイエンスフォーラム」(2016年3月17日)が国立国会図書館東京本館を会場に行われた<sup>(23)</sup>。その中で、研究データ問題を扱う国内の新たなコミュニティあるいは協議会のようなものを設置してはどうか、との提案があった。人が動いてこそ新たなものが生まれてくる。今後の我が国の活動に期待したい。

- (1) “RDA Seventh Plenary Meeting, Tokyo, Japan”. <https://rd-alliance.org/plenaries/rda-seventh-plenary-meeting-tokyo-japan/>, (accessed 2016-04-19). “データシェアリングシンポジウム”. <https://jipsti.jst.go.jp/rda/>, (参照 2016-04-19).
- (2) 内閣府. “科学技術基本計画”. <http://www8.cao.go.jp/cstp/kihonkeikaku/index5.html>, (参照 2016-05-20).
- (3) 例えば “OECD Principles and Guidelines for Access to Research Data from Public Funding”, OECD, 2007. <http://www.oecd.org/sti/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm>, (accessed 2016-04-19). “3. Open Scientific Research Data”. G8 Science Ministers Statement. GOV.UK, 2013-06-13. <https://www.gov.uk/government/news/g8-science-ministers-statement>, (accessed 2016-04-19). RDA Europe. “The Data Harvest”. [https://europe.rd-alliance.org/sites/default/files/repository/files/TheDataHarvestReport\\_%20Final.pdf](https://europe.rd-alliance.org/sites/default/files/repository/files/TheDataHarvestReport_%20Final.pdf), (accessed 2016-04-19).
- (4) 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 2016-05-20).
- (5) “研究データのオープンアクセス推進に関わる国際組織の動向”. オープンサイエンス推進に関するフォローアップ検討会 (第1回). [http://www8.cao.go.jp/cstp/tyousakai/opnsclfwp/lkai/1\\_followup\\_sankou4\\_7.pdf](http://www8.cao.go.jp/cstp/tyousakai/opnsclfwp/lkai/1_followup_sankou4_7.pdf), (参照 2016-05-20).
- (6) “オープンサイエンス推進に関するフォローアップ検討会”. 内閣府. <http://www8.cao.go.jp/cstp/tyousakai/opnsclfwp/index.html>, (参照 2016-05-20).
- (7) “Our Funders”. RDA. <https://rd-alliance.org/about-rda/our-funders.html>, (accessed 2016-05-20).
- (8) 恒松直幸, 加藤齊史, 大濱隆司, 村山泰啓, 研究データ同盟 (Research Data Alliance) 第2回総会. 情報管理, 2014, vol. 56, no. 10, p. 724-727. <http://doi.org/10.1241/johokanri.56.724>.
- (9) F. Berman, R. Wilkinson, and J. Wood. “Building Global Infrastructure for Data Sharing and Exchange Through the Research Data Alliance”. <http://doi.org/10.1045/january2014-berman>.
- (10) Hilary Hanahoe, 私信 (2016-05-17). 第7回総会では、中国8名、韓国4名、台湾3名、インド2名、パキスタン1名の参加があった。増加傾向なのは中国で、第1、3回総会で4~5名だったのが第6、7回総会では8名ずつ参加があった。
- (11) European Open Science Cloud. <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>, (accessed 2016-05-09).

本施策は、既存および将来のデータ基盤を融合してオープンな研究データ基盤を提供するとともに、これにより科学、ビジネス、公共サービスの加速、イノベーションの創出を図るイニシアチブで、欧州デジタル単一市場戦略 (Digital Single Market Strategy, 後述) の中に位置づけられる。2015年9月に高級専門検討会が組織され、ECから提案公募が開始された。

- (12)「研究データ基盤」は、RDA等で議論されているオープンサイエンスや分野横断的なデータ共有の文脈では、多くの技術的・社会的要素を持つ複合的なものと考えられる。例えば、情報通信技術 (高速インターネット網、計算機クラウドなど)、情報処理 (ビッグデータ・アナリティクス、データ・マイニング、統計処理など)、データの組織化・正準化と検索・相互利用 (識別子、メタデータ、レジストリ等)、データの信頼性・ポリシー・法的権利、各分野ごとの専門的差異などが検討課題としてあげられる。
- (13) EUDAT (European Data Infrastructure, 後出) や OpenAIRE などがあげられる。
- (14) Data Seal of Approval.  
<http://www.datasealofapproval.org/en/>, (accessed 2016-05-20).
- (15) 宮入暢子. 研究者識別子 ORCID: 活動状況と今後の展望. 情報管理, 2016, vol. 59, no. 1, p. 19-31.  
<http://doi.org/10.1241/johokanri.59.19>.
- (16) EUDAT.  
<https://www.eudat.eu/>, (accessed 2016-05-20).
- (17) EDISON.  
<http://edison-project.eu/>, (accessed 2016-05-20).
- (18) RDA-P7のサイドイベント (co-located events) は以下のようなものが行われた:  
データシェアリングシンポジウム, "RDA New Comers", "Publishing Data: Adoption and Implementation", "Competences and Skills for Data and Research Infrastructures", "Data Perspective beyond Alliances", "EUDAT/ROIS Collaborative Data Infrastructure Workshop", "Extend discussion for Library for Research Data IG and Long Tail of Research Data IG of RDA", "Beyond Compliance in Data Management", "Asian OA Community kickoff meeting".  
<https://rd-alliance.org/plenary-meetings/seventh-plenary/programme.html>, (accessed 2016-05-20).
- (19) "OECD Global Science Forum".  
<http://www.oecd.org/sti/sci-tech/oecdglobalscienceforum.htm>, (accessed 2015-05-20).
- (20) 例えば  
OECD. "Data-Driven Innovation".  
<http://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm>, (accessed 2016-05-10).  
European Commission - DG RTD. "Digital Single Market".  
[http://ec.europa.eu/growth/single-market/index\\_en.htm](http://ec.europa.eu/growth/single-market/index_en.htm), (accessed 2016-05-10).
- (21) Y Murayama. "Research Data Sharing and Frameworks".  
<http://www.slideshare.net/YasuhiroMurayama/icsti-annual-meeting-2014-tokyo>, (accessed 2016-04-19).
- (22) European Commission. Digital Single Market: Bringing down barriers to unlock online opportunities.  
<http://ec.europa.eu/priorities/digital-single-market/>, (accessed 2016-05-23).
- (23) "研究データとオープンサイエンスフォーラム～RDA 東京大会における議論を踏まえた研究データ共有の最新動向～". 国立国会図書館.  
<http://www.ndl.go.jp/jp/event/events/201603forum.html>, (参照 2016-05-20).

[受理: 2016-05-24]

Murayama Yasuhiro

International Framework of Research Data Sharing, and Cooperation with Librarians: A Report of 7th RDA Plenary Meeting