SPARC Japan

ISSN 2188-8248

国際学術情報流通基盤整備事業
（SPARC Japan）
年報

平成29（2017）年度

国立情報学研究所

SPARC Japan ロゴについて

表紙のロゴは、SPARC North America とあわせて 2017 年 6 月に新たに定めました。アスタリスクに似たシンボルは、デジタル環境における SPARC の確固とした立場を表現し、そのエネルギーとダイナミズムを強調しています。旧ロゴの新しいアイディアの誕生や知識のオープンシェアリングを拡大するといった”spark”のイメージを引き継ぎつつ、シンプルになりました。

目次

巻頭言 1
１　概要 2
１．１ 第５期（平成２８～３０年度）の活動概要 2
１．１．１ 第５期基本方針 2
１．１．２ 第５期事業計画 2
１．２ 平成２９年度活動 3
１．２．１ SPARC Japan セミナー 3
１．２．２ 海外動向調査 5
１．２．３ arXiv.org コンソーシアム事務局 5
１．２．４ SCOAP³ 支援 6
１．２．５ CLOCKSS 支援 6
１．２．６ 論文公表実態調査 6
１．２．７ 平成２８年度国際学術情報流通基盤整備事業(SPARC Japan)年報の発行 7
１．２．８ 高エネルギー物理学分野の情報サービスに係る国際連携協定への対応 7
２　委員会等開催記録 8
２．１ 国際学術情報流通基盤整備事業運営委員会 8
２．２ SPARC Japan セミナー企画 WG 8
３　委員名簿 8
３．１ 国際学術情報流通基盤整備事業運営委員会 8
３．２ SPARC Japan セミナー企画 WG 9
４　総合年表 10
５　刊行物一覧 22
５．１ 国際学術情報流通基盤整備事業(SPARC Japan)年報 22
５．２ SPARC Japan ニュースレター 22
５．３ SPARC Japan セミナードキュメント 22
６　資料　ニュースレター再掲 25

巻頭言

2003 年度から始まった本事業も、2017 年度で第 5 期の二年目を迎えました。一年間の活動をとりまとめて、ここに報告いたします。

第 5 期は、国際的なオープンアクセス（OA）イニシアティブとの協調、そして OA やオープンサイエンスの国内アドボカシーに注力しております。ご存知のように、今や研究活動は国際的な連携を抜きには語れません。更に近年では、学術雑誌の購読価格高騰に端を発した OA、そして科学研究の効率化や科学の発展のみならず，イノベーションの喚起や研究公正を背景としたオープンサイエンスの活動が、日本の研究活動にも深く関係してくるようになりました。このような情勢を踏まえて、学術情報の流通を如何に推進し、サステイナブルな仕組みを作るかが大きな課題となっています。

具体的には、2017 年度においては、arXiv.org や SCOAP³ といった国際的な OA イニシアティブとの協調を進めました。arXiv.org については昨年度に引き続き、Member Advisory Board に日本コンソーシアム代表として京都大学の引原隆士図書館機構長にご出席いただき、機関リポジトリ等との連携について提案、積極的にご議論いただきました。SCOAP³ は高エネルギー物理学分野の査読付き雑誌論文の OA 化を目的とした国際連携プロジェクトですが、2018 年からアメリカ物理学会（APS）刊行誌の一部が同プロジェクトの対象となることから、国内参加機関のとりまとめの役割を担うと共に、日本物理学会誌に啓発広告を掲載しております。更に、欧州原子核研究機構（CERN）が提供しているデータベース「INSPIRE」に収載された高エネルギー物理学分野の論文情報の品質向上のため、弊所の実務研修制度を利用し、大学図書館員 1 名を派遣、CERN においてキュレーションの経験を積んでいただきました。

アドボカシー活動としては、「オープンアクセス温故知新-ふりかえって次をみつける-」を年間テーマとして、SPARC Japan セミナーを 3 回開催しました。第 1 回セミナーでは、図書館員と研究者が研究データをいかに管理・流通させていくかという視点を共有し、そこからもたらされる両者の新たな関係について、第 2 回セミナーでは、プレプリントサーバの機能や、その変遷を振り返りつつ、研究推進への寄与や持続性のあるビジネスモデルの構築、質の確保といった将来の課題を見据えて、これからのオープンアクセスを、第 3 回セミナーでは、社会全般にわたるデジタル化や学術コミュニケーションの形態の変化の中で、科学と学術の本来の姿を議論することで、学術コミュニティが向かうべき方向性を見出すことを目的に、各回、講演とパネルディスカッションを行いました。なお、第 3 回セミナーでは、SPARC North America のエグゼクティブディレクターを務める Heather Joseph 氏をお招きし講演いただいております。また、新たな試みとして、セミナーの動画中継を始めました。有難いことにご好評いただいており、今後も継続したいと考えております。

こうした諸活動は、大学図書館コンソーシアム連合（JUSTICE）やオープンアクセスリポジトリ推進協会（JPCOAR）はもとより、研究者コミュニティ等と協調しつつ、継続的に行っていくことが重要です。私自身は、この 3 月末を以て委員長を退任することになります。この場をお借りして、この事業を支えてくださった方々、特にセミナーを企画した企画 WG の委員の方々のご尽力に感謝の念を表したいと存じます。

2018 年 3 月 31 日
国際学術情報流通基盤整備事業運営委員会 委員長
安達 淳

１ 概要

１．１ 第５期（平成２８～３０年度）の活動概要

１．１．１ 第５期基本方針

第５期においても、第４期の活動を継承し、国内外の OA イニシアティブや関係組織と連携し、オープンアクセス等を推進し、学術情報流通の更なる発展に取り組むことを基本方針とする。

特に米国 SPARC と連携し、日本のオープンアクセス活動を国際的に発信する。オープンアクセス等の推進にあたっては、まずその課題を把握することに努めると共に、「大学図書館と国立情報学研究所との連携・協力推進会議」の下の機関リポジトリ推進委員会および大学図書館コンソーシアム連合等との協調を一層強化し、学術情報流通の発展に向けて参加意識を強める方向でアドボカシー活動を継続的に行っていく。

１．１．２ 第５期事業計画

SPARC Japan　第５期の事業は次の 4 つを柱として計画することが、平成 27 年度第 3 回国際学術情報流通基盤整備事業運営委員会で決定した。

⑴国際的な OA イニシアティブとの協調

国際イニシアティブに参画し、日本の窓口としての役割を果たすとともに、その活動・成果のアピールに努める。これらも含めて、国際的な動向を注視し、必要な対応を行う。

⑵学術情報流通にかかわるアドボカシー活動

「大学図書館と国立情報学研究所との連携・協力推進会議」等の組織と連携しつつ、オープンアクセスやオープンサイエンス、学協会出版の国際流通に係るアドボカシー活動を継続して実施する。

⑶オープンサイエンスへの活動スコープの拡大

研究成果のオープンアクセス、イノベーションの基盤となる可能性を秘めたオープンデータ、加えて高等教育の基本的構成要素の再考を迫るオープンエデュケーションなどへの関心の高まりにあわせて、理工学分野だけではなく、人文科学・社会科学分野の動向等に関して適時の情報提供を実現する。また、大学図書館におけるオープンサイエンスの取組み、研究データの管理等への関与について、戦略的な検討を行う。

⑷オープンアクセスに関する基礎的情報の把握

第４期に引き続き、オープンアクセスに関する基礎的情報を把握するために実態調査等を行う。各大学・研究機関の研究戦略を考える上で、データを集め分析するために、図書館が一定の役割を果たすことも検討する。

１．２ 平成２９年度活動

１．１．２の事業計画のもと、平成 29 年度は次のプロジェクトを実施した。

１．２．１ SPARC Japan セミナー

アドボカシー活動として、SPARC Japan セミナーを 3 回実施した。セミナー企画ワーキンググループ（以下、「WG」という。）を立ち上げ、WG メンバー全員で年度を通した全体テーマ、セミナー各回のテーマの割り振りを検討し、各回に担当者を置いて企画・実施した。セミナー終了後、SPARC Japan ニュースレター（以下、「NL」という。）を発行・ウェブ配信した。

SPARC Japan セミナーの記録

年間テーマ　「オープンアクセス温故知新-ふりかえって次をみつける-」

第 1 回	平成 29 年 9 月 13 日（水）13:00～17:20　（NII 19 階会議室） 「図書館員と研究者の新たな関係：研究データの管理と流通から考える」 https://www.nii.ac.jp/sparc/event/2017/20170913.html 〔司会〕能勢　正仁　（京都大学大学院理学研究科） ・「開会挨拶/趣旨説明」 能勢　正仁　（京都大学大学院理学研究科） ・「研究者にとってのデータの意味と大学におけるデータ管理への期待」 倉田　敬子　（慶應義塾大学文学部） ・「学術リポジトリは研究者と図書館員を繋げるのか？」 大澤　剛士　（農業・食品産業技術総合研究機構農業環境変動研究センター） ・「研究データ管理の組織的支援と図書館の役割について」 西薗　由依　（鹿児島大学/JPCOAR 研究データタスクフォース） ・「新たな学術情報流通において JPCOAR スキーマが果たす役割」 片岡　朋子　（お茶の水女子大学/JPCOAR メタデータ普及タスクフォース） ・「研究データ利活用協議会(RDUF)紹介」 武田　英明　（国立情報学研究所/研究データ利活用協議会） ・「全体議論」 〔モデレーター〕能勢　正仁（京都大学大学院理学研究科） 企画 WG（五十音順、◎は主査） 中谷 昇　（鳥取大学学術情報部） 能勢　正仁　（京都大学大学院理学研究科）◎ 林 賢紀　（国際農林水産業研究センター）
	参加人数：60 名　（定員：60 名） 動画中継利用件数※：286、アーカイブ動画利用件数※：312 NL 第 33 号　（2017 年 12 月）　（参照：7　資料　ニュースレター再掲）
第 2 回	平成 29 年 10 月 30 日（月）11:00～16:40　（NII 12 階会議室） 「プレプリントとオープンアクセス」 https://www.nii.ac.jp/sparc/event/2017/20171030.html 〔司会〕坊農　秀雅　（情報・システム研究機構ライフサイエンス統合データベースセンター） ・「開会挨拶/趣旨説明」 坊農　秀雅　（情報・システム研究機構ライフサイエンス統合データベースセンター） ・「arXiv.org の次世代システムの公開と戦略」 引原　隆士　（京都大学図書館機構長/arXiv.org 会員コンソーシアム代表） ・「学術情報共有とオープンアクセスの未来」 Gregg Gordon　（Managing Director, SSRN） ・「化学分野におけるプレプリントの位置付け・課題等について」

	生長　幸之助　（東京大学大学院薬学系研究科/化学ポータルサイト Chem-Station 副代表） ・「生命科学分野におけるプレプリントの位置付けや経験について，統合 TV について」 小野　浩雅　（情報・システム研究機構ライフサイエンス統合データベースセンター） ・「全体議論」 〔モデレーター〕坊農　秀雅　（情報・システム研究機構ライフサイエンス統合データベースセンター） 企画 WG（五十音順、◎は主査） 梶原　茂寿　（室蘭工業大学附属図書館） 笹渕　洋子　（早稲田大学図書館） 坊農　秀雅　（情報・システム研究機構ライフサイエンス統合データベースセンター）◎
	参加人数：63 名　（定員：70 名） 動画中継利用件数※：190、アーカイブ動画利用件数※：488 NL 第 34 号（2018 年 3 月）　（参照：7　資料　ニュースレター再掲）
第 3 回	平成 30 年 2 月 21 日（水）10:30～17:00　（NII 12 階会議室） 「オープンサイエンスを超えて」 https://www.nii.ac.jp/sparc/event/2017/20180221.html 〔司会〕林　和弘（科学技術・学術政策研究所） ・「趣旨説明」 蔵川　圭　（国立情報学研究所） ・「オープンサイエンスを真に理解する：その有益性の潜在能力，脆弱性，機能的パフォーマンスの問題，これらの解決策を講じない方法」 Paul A. David　（Stanford University） ・「データ駆動型の科学研究エコシステムとしてのオープンサイエンスー過去のコミュニティ実践事例と日本の視点」 村山　泰啓（情報通信研究機構/ICSU-World Data System） ・「オープンリサーチを可能にするには」 Heather Joseph　（SPARC North America） ・「デジタル時代の研究プロセスと大学，大学図書館における支援のあり方」 倉田　敬子　（慶應義塾大学文学部） ・「デジタル化時代の研究者のために図書館が構築すべき学術情報環境」 市古　みどり　（慶應義塾大学三田メディアセンター） ・「パネルディスカッション」 「討論への問題提起：オープンサイエンスの知識論を考える」 〔モデレーター〕深貝　保則（横浜国立大学大学院　国際社会科学研究院） 企画 WG（五十音順、◎は主査） 梶原　茂寿　（室蘭工業大学附属図書館） 蔵川　圭　　（国立情報学研究所）◎ 林　和弘　　（科学技術・学術政策研究所）
	参加人数：67 名（定員：70 名） 動画中継利用件数※：205、アーカイブ動画利用件数※：動画未公開 NL 第 35 号（2018 年 10 月）　（参照：7　資料　ニュースレター再掲）
年間	参加人数：190 名　（定員計 200 名、平均 63 名） 動画中継利用件数※：681、アーカイブ動画利用件数※：800

※2018 年 3 月 11 日時点

１．２．２ 海外動向調査

下記の国際会議等に参加し、情報収集を行った。

・RDA (Research Data Alliance) 9th Plenary Meeting
（4 月 5-7 日 Barcelona、Spain）に NII 武田教授，山地准教授，船守准教授，金澤准教授，蔵川特任准教授を派遣した。

・COAR (Confederation of Open Access Repository) Annual meeting 2017
（5 月 8-10 日 Venice、Italy）に NII 山地准教授を派遣した。

・IDF Strategic Meeting International DOI Foundation
（6 月 13-14 日 Daejeon、Korea）に NII 武田教授を派遣した。

・DOI Strategy Meeting
（6 月 15 日 Seoul、Korea）に NII 武田教授を派遣した。

・OR2017 (The 12th Annual International Conference on Open Repositories)
（6 月 27-30 日 Brisbane、Australia）にオープンアクセスリポジトリ推進協会作業部会委員の大学図書館員 1 名を派遣した。

・RDA (Research Data Alliance) 10th Plenary Meeting
（9 月 18-21 日 Montreal、Canada）に NII 武田教授を派遣した。

・PIDapalooza (PID：Persistent ID)
（1 月 23-24 日 Girona、Spain）に NII 武田教授を派遣した。

・IDF Annual Meeting
（1 月 25-26 日 Barcelona、Spain）に NII 武田教授を派遣した。

１．２．３ arXiv.org コンソーシアム事務局

arXiv.org は物理学、数学、コンピュータサイエンス及び関連分野のプレプリントサーバで、コーネル大学図書館が運用している。2017 年には電気工学システム科学、経済学が新たに加わった。2014 年 12 月に 100 万論文を突破し、2017 年には約 137 万件、新規登録数は年間 12 万件、ダウンロード数は年間約 1.8 億件以上で、総ダウンロード数は 10 億件を超えた。利用件数上位の機関による財政支援があり、2013 年に開始した「arXiv 会員制プログラム」により 2017 年末現在で 27 カ国 221 機関が参加している。

日本においては NII が各大学の意思確認を取りまとめて支援してきた。2014 年 4 月にコーネル大学から、日本の会員館でコンソーシアムとして参加することについて打診があり、会員に確認の後、コンソーシアム契約に切り替えを行った。会費はコンソーシアム価格のため 10%減となった。2017 年は利用件数上位 300 位までの大学に会員申請の意向調査を行った結果、2017 年 12 月末現在の会員数は 14 機関である。

さらに 2015 年度には、本コンソーシアム名を日本研究図書館コンソーシアム（英文名：Consortium of Japanese Research Libraries: Coordinated by National Institute of Informatics (NII)、英文略称： NII Japan Consortia）とした。また、引原隆士 京都大学図書館機構長が本コンソーシアムの代表に就任し、2016 年から arXiv.org の Member Advisory Board（MAB）に参加し、2016 年に続き、2017 年は 10 月 5 日開催の会議に出

席した。

１．２．４ SCOAP³支援

2014 年から開始した SCOAP³について 2017 年も参加意向および連絡先を確認し、日本の大学図書館からの拠出金を、日本のナショナル・コンタクト・ポイントである NII がとりまとめて支払った。なお、フェーズ 1（2014 年～2016 年）における日本の参加機関は 34 機関で、フェーズ 2（2017 年～2019 年）は 2017 年に 40 機関、さらに 2018 年からアメリカ物理学会が対象誌として加わったことをうけ、2018 年 3 月末現在で 67 機関が参加を表明している。

SCOAP³により OA 化された論文は 17,850 件に達している（2014 年～2017 年)。論文あたりのコストを算出すると、1,032Euro であり、一般的な Gold OA 誌の APC に比べて低く抑えられている。

SCOAP³の OA 論文を収載した SCOAP³リポジトリとその API が既に公開されているが、リポジトリは DOI 付与、CC BY ライセンス表示、XML 形式で公開されており、テキストマイニング、データマイニングが可能である。

2018 年からはアメリカ物理学会（American Physical Society：APS）が SCOAP³に加わり、APS の刊行する Physical Review C、Physical Review D、Physical Review Letters の 3 誌に掲載されている HEP 分野の論文がオープンアクセスになることが発表された。これにより、2018 年以降は世界中の HEP 分野の論文の約 90%がオープンアクセスとなった。2018 年，2019 年の各購読機関の APS 契約金額からは SCOAP³拠出額が削減される。

SCOAP³評議会日本代表委員である NII 安達教授が Executive Meeting（9 月 26 日、テレビ会議）及び Forum（2017 年 12 月 7 日、Webinar）に出席した。また、さらに多くの機関に参加を求めるため、本年度は物理学研究者向けに動画を作成・公開するとともに、日本物理学会誌に広告を掲載するといった広報活動を行った。

１．２．５ CLOCKSS 支援

CLOCKSS（Controlled Lots of Copies Keep Stuff Safe）は、全世界の研究者のためにデジタル資源（Web ベースの学術文献等）の長期保存を実現することを目的とし、アーカイブとそれを運営するコミュニティを構築して、コンテンツが出版社から提供されなくなった場合にアーカイブ上のコンテンツを広く利用できるようにするなどの取り組みを行っている。

日本においては 2013 年から NII が各大学の参加意向確認および年会費支払いのとりまとめを行っている。なお、日本からは 2017 年 4 月末現在で 98 機関が参加している。

１．２．６ 論文公表実態調査

2015 年度に、大学図書館コンソーシアム連合（JUSTICE）と連携し、JUSTICE の下に論文公表実態調査チームを設け、我が国における論文公表と APC（Article Processing

Charge）の実態調査を開始した。平成 29 年度も引き続きフォローアップを行った。

１．２．７　平成２８年度国際学術情報流通基盤整備事業(SPARC Japan)年報の発行

平成 28 年度の活動状況をまとめ、平成 30 年 3 月に発行した。また、平成 27 年度 SPARC Japan 年報について、平成 30 年 3 月に英語版を発行した。

１．２．８　高エネルギー物理学分野の情報サービスに係る国際連携協定への対応

高エネルギー物理学分野に関する国際連携協定を欧州原子核研究機構（CERN)、高エネルギー加速器研究機構（KEK）及び国立情報学研究所（NII）の 3 機関が締結した。

CERN が提供する高エネルギー物理学分野の文献情報サービスである INSPIRE データベースを中心にデータキュレーションの実務に関する共同事業を行うために、平成 29 年 10 月から約 3 か月間、大学図書館員 1 名を CERN に派遣するとともに、双方の関係者によるテレビ会議を複数回行った。

２ 委員会等開催記録

２．１ 国際学術情報流通基盤整備事業運営委員会

開催日	議題
第 1 回 平成 29 年 9 月 26 日	1. 前回議事要旨（案）について 2. 平成 29 年度 SPARC Japan 活動状況について【報告】 3. 国際連携の状況について【報告・審議】 4. SPARC Japan の今後の活動方針について【審議】 5. JUSTICE における Open Access に係る活動状況報告【報告】 6. その他
第 2 回 平成 30 年 3 月 14 日	1. 前回議事要旨（案）について 2. 平成 29 年度 SPARC Japan 事業報告【報告】 3. JUSTICE における Open Access に係る活動状況報告【報告】 4. 国際連携の状況について【報告】 5. 平成 30 年度 SPARC Japan 活動計画について【審議】 6. 平成 30 年度 SPARC Japan セミナー企画ワーキンググループの設置について【審議】 7. SPARC Japan の今後の活動方針について【審議】 8. その他

２．２ SPARC Japan セミナー企画 WG

開催日	議題
キックオフミーティング 平成 29 年 6 月 21 日	1. 趣旨説明 2. 年間テーマ確定 3. スケジュール・分担検討 4. その他

３ 委員名簿

３．１ 国際学術情報流通基盤整備事業運営委員会

氏名	所属・役職	備考
逸村　裕	筑波大学 図書館情報メディア系 教授	1 号委員 （研究教育職員）
今井　浩	東京大学大学院 情報理工学系研究科 教授	1 号委員 （研究教育職員）
深貝　保則	横浜国立大学大学院国際社会科学研究院 教授	1 号委員 （研究教育職員）
倉田　敬子	慶應義塾大学 文学部 教授	1 号委員 （研究教育職員）
野﨑　光昭	高エネルギー加速器研究機構 教授 国際連携推進室 室長	1 号委員 （研究教育職員）
土屋　俊	独立行政法人 大学改革支援・学位授与機構 教授	1 号委員 （研究教育職員）

島　　文子	北海道大学附属図書館 事務部長 (国大図協 OA 会長館)	２号委員 （大学図書館関係者）
高橋　　努	東京大学附属図書館 事務部長	２号委員 （大学図書館関係者）
荘司　雅之	早稲田大学図書館 事務部長	２号委員 （大学図書館関係者）
市古　みどり	慶應義塾大学三田メディアセンター 事務長 (大学図書館コンソーシアム連合運営委員会委員長)	２号委員 （大学図書館関係者）
林　和弘	科学技術・学術政策研究所 科学技術予測センター 上席研究官	３号委員 （学会の関係者）
安達　淳	国立情報学研究所 副所長	１号委員 （研究教育職員） 委員長
武田　英明	国立情報学研究所 情報学プリンシプル研究系 教授 (JaLC 運営委員会委員長)	１号委員 （研究教育職員）
江川　和子	国立情報学研究所 学術基盤推進部 次長	２号委員 （大学図書館関係者）

３．２ SPARC Japan セミナー企画 WG

氏名	所属・役職
梶原　茂寿	室蘭工業大学 総務広報課 図書学術情報室長 （国立情報学研究所推薦）
能勢　正仁	京都大学大学院理学研究科附属地磁気世界資料解析センター 助教
中谷　昇	鳥取大学 学術情報部図書館情報課 （オープンアクセスリポジトリ推進協会推薦）
笹渕　洋子	早稲田大学図書館 調査役（電子資料担当） （大学図書館コンソーシアム連合運営委員会推薦）
林　和弘	科学技術・学術政策研究所 科学技術予測センター 上席研究官
林　賢紀	国立研究開発法人国際農林水産業研究センター 企画連携部情報広報室情報管理科情報管理係 （これからの学術情報システム構築検討委員会推薦）
坊農　秀雅	情報・システム研究機構 データサイエンス共同利用基盤施設 ライフサイエンス統合データベースセンター (DBCLS) 特任准教授
蔵川　圭	国立情報学研究所　学術コンテンツ課 特任准教授

4 総合年表

年度	評議会　運営委員会	主催イベント	その他のイベント
平成 15 （2003）	06/25　第 1 回評議会 07/14　事業参画提案の募集開始 08/01　第 1 回運営委員会 09/11　第 2 回運営委員会 09/17　第 2 回評議会（事業参画提案決定） 09/17　記者発表 10/08　作業グループ合同会議 03/22　第 3 回運営委員会 03/23　第 3 回評議会	07/02　学協会向け事業説明会（於：日本教育会館） 08/19　事業説明会（於：東北大学　東北大学附属図書館との共催） 01/21-29 Project Euclid 説明会（於：学術総合センター、東北大学、京都大学、名古屋大学） 02/23　SPARC/JAPAN 懇談会：参加学会への成果報告、新雑誌創刊構想説明（於：学術総合センター） 03/11　SPARC/JAPAN セミナー：生物系学協会誌をめぐる学術情報流通体制の将来　-UniBio Press のめざすもの-（於：東京大学附属図書館）	11/05　第 5 回図書館総合展フォーラム「SPARC／JAPAN：日本の国際学術コミュニケーションの変革」開催（於：東京国際フォーラム　国立大学図書館協議会・私立大学図書館協会主催） 11/20　国立大学図書館協議会電子ジャーナルタスクフォース（生物系、物理系、医学系の購読交渉）
平成 16 （2004）	05/28　第 1 回運営委員会 06/02　第 1 回評議会 06/07　参画提案募集開始 09/15　第 2 回運営委員会 09/22　第 2 回評議会（事業参画提案選定）	 07/07　学協会向け事業説明会（於：学術総合センター）	07/01　国立大学図書館協会総会ワークショップ：「国際学術情報流通基盤整備事業の活動」（於：大阪大学コンベンションセンター）

	10/14　作業グループ合同会議 03/07　第 3 回運営委員会 03/10　第 3 回評議会	09/27　Project Euclid 懇談会（Project Euclid への参画に関する技術的打ち合わせ、DPubS についての説明） 10/15　シンポジウム：学会出版と学術コミュニケーション活動の変革～SPARC/JAPAN を事例として～（於：広島大学中央図書館　広島大学図書館、国立情報学研究所、国立大学図書館協会国際学術コミュニケーション委員会共催） 10/19　緊急シンポジュウム「どうする日本の学術誌！」（於：早稲田大学総合学術情報センター　（社）高分子学会、（社）電子情報通信学会、東北数学雑誌編集委員会、（社）日本機械学会、（社）日本金属学会、（社）日本動物学会、（社）日本分析化学会、日本哺乳動物卵子学会、日本哺乳類学会、国立情報学研究所共催） 11/05　OUP 懇談会「Open Access の現状について」 11/25　第 6 回図書館総合展フォーラム「学術コミュニケーションの最先端：オープン・アクセスとセルフアーカイブ」（於：パシフィコ横浜） 01/27　ワークショップ「電子ジャーナルのビジネスモデル構築と学術出版をめぐる動向」（於：日本教育会館） 03/24　シンポジウム「SPARC の現状と課題：学術雑誌・機関レポジトリ・オープン・アクセス」（於：早稲田大学）	10/19-20 Project Euclid DPubS Conference に参加（於：コーネル大学）
平成 17 （2005）	06/06　第 1 回運営委員会 06/08　第 1 回評議会	05/19　SPARC/JAPAN 連続セミナー第 1 回「Nature の歴史、今、未来を語る－Nature の編集方針」 06/29　SPARC/JAPAN 連続セミナー第 2 回「電子投稿査読システムとは何か－今、日本で使えるシステム」JST「J-STAGE 投稿審査システム」 07/09-10 電子ジャーナル利用の現在と未来に関するクローズド・ワークショップ（於：経団連ゲストハウス、静岡） 07/15　SPARC/JAPAN 連続セミナー第 3 回「オープン・アクセスの理念と実践－研究者・図書館・学術誌」	06/21-22 JISC International Solutions for the Dissemination of Research に出席、討議（ロンドン） 07/07-08 エルゼビア・ライブラリ・コネクト・セミナー2005「ユーザーを理解する（Understanding Users）」（於：京都・東京、エルゼビア・ジャパン主催、NII 後援）

	10/13　第 2 回運営委員会 10/26　第 2 回評議会（事業参画提案選定）	07/20　UniBio Press の挑戦 － 学会の新しいビジネスモデル（於：茨城大学　茨城大学図書館主催） 09/22　SPARC/JAPAN 連続セミナー第 4 回「電子ジャーナルをどう作成し、どう公開するかー学協会、企業の試み」 10/06　SPARC/JAPAN 連続セミナー第 5 回「主体である研究者は何をすべきかー電子ジャーナル時代を迎えて」（於：つくば国際会議場社団法人日本動物学会第 76 回大会関連シンポジウムとの共催） 11/24　SPARC/JAPAN 連続セミナー臨時回「Journal of Bioscience and Bioengineering WEB 投稿審査システム」説明会・デモンストレーション 11/30　SPARC/JAPAN 連続セミナー第 6 回「第 7 回図書館総合展フォーラム COUNTER プロジェクト：オンライン利用統計の国際標準について」（於：パシフィコ横浜） 12/01　COUNTER プロジェクトに関するクローズド・ワークショップ 12/12　SPARC/JAPAN 連続セミナー第 7 回「日本の学術誌における英文校閲を考える」 01/31　SPARC/JAPAN 連続セミナー第 8 回「学術情報流通をめぐる最近の動向と技術標準：Google Scholar、 CrossRef、 OAI-PMH、 etc.」 02/10　SPARC/JAPAN 連続セミナー第 9 回「SPARC/JAPAN 選定誌によるラップアップセッション」	09/15　山口大学図書館セミナー2005「日本の電子ジャーナルの現況」学術コミュニケーションの今日：SPARC/JAPAN の挑戦（於：山口大学　山口大学学術情報機構図書館主催） 09/16　京都大学学術情報・電子ジャーナルシンポジウム「「大学における学術情報資源の整備―電子ジャーナル時代の学術コミュニケーションの変革―」（於：京都大学　京都大学附属図書館と NII の共催） 12/09　長崎大学附属図書館連続講演会第二回講演会「学術情報発信の新しい動向」：SPARC/JAPAN の活動と課題（於：長崎大学　長崎大学附属図書館主催）
	02/15　第 3 回運営委員会 02/24　第 3 回評議会		
平成 18 （2006）		06/30　SPARC Japan 連続セミナー2006 第 1 回「海外商業出版社から見た日本の学術コミュニティ」 07/26　SPARC Japan 連続セミナー2006 第 2 回「e-Journal の販促とライセンシング：海外の状況と海外市場における日本ジャーナルの展望」	03　米国研究図書館協会（ARL）と MOU を締結 07/03-04 エルゼビア・ライブラリ・コネクト・セミナー2006「From “ Search ” to “ Find ” ～ 必要な情報を見つけやすい環境づくり ～」（於：東京・大阪、エルゼビア・ジャパン主催、NII 後援）

	09/08　第 1 回運営委員会 01/30　第 2 回運営委員会	09/05　Sally Morris 氏講演会「Introducing ALPSP」 09/29　SPARC Japan 連続セミナー2006 第 3 回「Web 投稿審査システムの検証：ビフォーアフター」 11/02　SPARC Japan 連続セミナー2006 第 4 回「大学図書館から学術出版社への要望：COUNTER を例にして」 11/20　第 8 回図書館総合展フォーラム「TRANSFER－出版社間のジャーナル移行に伴う問題点とその解決に向けて」（於：パシフィコ横浜） 12/14　SPARC Japan 連続セミナー2006 第 5 回「著作権：学会の権利、著者の権利、機関リポジトリへの対応」 12/18-19　「デジタル巨人の肩の上に立つ」機関リポジトリ、e-サイエンス、および学術コミュニケーションの将来に関する国際シンポジウム（於：都市センターホール） 01/30　SPARC Japan 連続セミナー2006 第 6 回「e-Journal の販売とライセンシング(2)- 販売のプロに学ぶ成功の秘訣」 03/05　SPARC Japan 連続セミナー2006 第 7 回「計量書誌学からジャーナル・論文のパフォーマンスを測る」	
平成 19 （2007）	06/12　パートナー誌合同会議 07/19　第 1 回運営委員会	07/17　SPARC Japan 連続セミナー2007 第 1 回「計量書誌学からジャーナル・論文のパフォーマンスを測る-2-」 10/02　SPARC Japan 連続セミナー2007 第 2 回「Web 投稿審査システムの検証パート 3 稿より良いシステムを目指して―」	05/15　UniBio Press セミナー「生物系ジャーナル の挑戦－より明確に、より広く、その情報を伝えるために」（於：学術総合センター　UniBio Press 主催） 05/17　UniBio Press セミナー「生物系ジャーナル の挑戦－より明確に、より広く、その情報を伝えるために」（於：京都大学附属図書館　UniBio Press 主催） 08/05-11 41th IUPAC （International Union of Pure and Applied Chemistry）化学会議出展（於：トリノ） 08/20-22　234th ACS 秋季大会出展（於：ボストン）

	12/14　パートナー誌と大学図書館の合同会議「SPARC Japan パートナー誌のコンソーシア購入に向けて」 02/29　第 2 回運営委員会	11/02　SPARC Japan 連続セミナー2007 第 3 回「メタデータ Publishing の現在―電子ジャーナル主体の製作・出版に必要なもの」 11/09　第 9 回図書館総合展プレゼンテーション「日本の英文トップ電子ジャーナルの挑戦―図書館総合展プレゼンテーパートナー誌からの提案―」（於：パシフィコ横浜） 01/17　SPARC Japan-ALPSP 特別セミナー（第 4 回 SPARC Japan 連続セミナー2007）「学術出版と学会 Journal Publishing and Scholarly Societies」 01/18　ALPSP トレーニングコース「Introduction to Journal Publishing」	11/07-09 第 9 回図書館総合展出展（於：パシフィコ横浜）
平成 20 （2008）		04/22　SPARC Japan セミナー2008 第 1 回「研究成果発表の手段としての学術誌の将来」 06/24　SPARC Japan セミナー2008 第 2 回「学術出版と XML 対応-日本の課題」 07/10　SPARC Japan セミナー2008 第 3 回「韓国コンソーシアム事情 - 海外展開を目指して -」 09/02-03 RIMS 研究集会（第 4 回 SPARC Japan セミナー2008）「紀要の電子化と周辺の話題」（於：京都大学数理解析研究所 京都大学数理解析研究所主催） 10/14　SPARC Japan セミナー2008 （Open Access Day 特別セミナー）「日本における最適なオープン・アクセスとは何か？」	06/15-17 SLA（Special Libraries Association 米国専門図書館協会）年次総会出展（於：シアトル） 06/26　第 55 回国立大学図書館協会総会出展（於：東北大学） 07/13-15 中国化学会学術年会出展（於：天津） 08/17-19 236th ACS National Meeting & Exposition 出展（於：フィラデルフィア） 09/11-12 私立大学図書館協会総会出展（於：國學院大學） 09/16-20 2nd EuCheMS Chemistry Congress 出展（於：トリノ） 09/25-26 KESLI（Korean Electronic Site License Initiative）電子情報 EXPO での発表、出展（於：大田） 10/12-15 15th North American ISSX (International Society for the Study of Xenobiotics)) Meeting での広報（於：サンディエゴ） 10/27-30 ISAP2008（International Symposium on Antennas and Propagation）出展（於：台湾）

			11/13-14 INFOPRO2008 プロダクトレビュー参加・出展（於：日本科学未来館）
		11/17-18 SPARC Digital Repositories Meeting 2008（於：ボルチモア SPARC、SPARC Europe、 SPARC Japan 共同主催） 11/25 SPARC Japan セミナー2008 第 6 回「IF を越えて-さらなる研究評価の在り方を考える」 11/27 SPARC Japan セミナー2008 第 7 回（第 10 回図書館総合展・学術情報オープンサミット 2008 フォーラム）「Open Access Update」 12/16 SPARC Japan セミナー2008 第 8 回「日本で使える電子ジャーナルプラットフォーム」	
	12/24 第 1 回運営委員会	01/22-26 Project Euclid と数学系ジャーナルの打ち合せ（於：国立情報学研究所、京都大学、東京工業大学） 02/13 SPARC Japan セミナー2008 第 9 回「SPARC Japan 選定誌がやってきたこと」	12/17-20 EUC2008（International Conference On Embedded and Ubiquitous Computing）出展（於：上海）
	03/10 第 2 回運営委員会 03/27 パートナー誌合同会議 03/27 第 3 回運営委員会		03/16-20 APS March Meeting 2009（米国物理学会年会）出展（於：ピッツバーグ）
平成 21 （2009）	10/05 第 1 回運営委員会	06/25 SPARC Japan セミナー2009 第 1 回「研究者は発信する－多様な情報手段を用い、社会への拡がりを求めて」 08/04 SPARC Japan セミナー2009 第 2 回「非営利出版のサステイナビリティとは－OUP に学ぶ」 09/08-09 RIMS 研究集会（第 3 回 SPARC Japan セミナー2009）「数学におけるデジタルライブラリー構築へ向けて－研究分野間の協調のもとに」 09/17 日本動物学会大会（第 4 回 SPARC Japan セミナー2009）「ZS プロジェクトについて」 10/20 Open Access Week （第 5 回 SPARC Japan セミナー2009）「オープンアクセスのビジネスモデルと研究者の実際」 11/11 第 6 回 SPARC Japan セミナー2009 （第 11 回図書館総合展学術情報オープンサミット 2009 フォーラム）「NIH Public Access Policy とは何か」	11/25 第 9 回アジア太平洋生物化学工学会議（APBioChEC 2009）に SPARC Japan の化学系パートナー誌が出展（於：神戸国際会議場）

	03/23　第 2 回運営委員会	12/11　第 7 回 SPARC Japan セミナー2009「人文系学術誌の現状ー機関リポジトリ、著作権、電子ジャーナル」 02/02　第 8 回 SPARC Japan セミナー2009「Marketing to Libraries Worldwide」 02/03　ALPSP トレーニングコース「Effective Journals Marketing」	12/03-04 DRFIC 2009　デジタルリポジトリ連合国際会議 2009（於：東京工業大学　DRF（デジタルリポジトリ連合）と NII の共催）
平成 22（2010）		06/23　第 1 回 SPARC Japan セミナー2010「学会の仕事とその経営を知る」 07/06　第 2 回 SPARC Japan セミナー2010「ジャーナル出版―海外学会の現状」 08/24　第 3 回 SPARC Japan セミナー2010「図書館の仕事を知る-学術雑誌の購読と利用-」 09/16　第 4 回 SPARC Japan セミナー2010（RIMS 研究集会）「数学におけるデジタルライブラリー構築へ向けて」 09/24　第 5 回 SPARC Japan セミナー2010（社団法人 日本動物学会 第 81 回大会）「日本の学術情報流通　10 年後を見据えて」 10/20　第 6 回 SPARC Japan セミナー2010 Open Access Week「日本発オープンアクセス」 11/08- 09 SPARC Digital Repositories Meeting (デジタルリポジトリ会議）（於：ボルチモア SPARC、SPARC Europe、 SPARC	08/19　International Congress of Mathematicians (国際数学者会議) に出展　（於：ハイデラバード） 08/22- 26 American Chemical Society (ACS) 2010 年 秋季大会に出展　（於：マサチューセッツ） 08/29-09/02 3rd EuCheMS Chemistry Congress（第 3 回ヨーロッパ化学会議）に出展　（於：ニュルンベルク）

	03/16　第 1 回運営委員会	Japan 共催） 12/10　シンポジウム 「大学からの研究成果オープンアクセス化方針を考える」（於：東京大学　国立大学図書館協会と NII の共催） 01/14　第 7 回 SPARC Japan セミナー2010 「著者 ID の動向」 02/03　第 8 回 SPARC Japan セミナー2010 「世界における”日本の論文／日本の学術誌”のインパクト」 03/08　TIB（ドイツ技術情報図書館）/ ZB MED（ドイツ医学中央図書館）/ NII（国立情報学研究所）MoU 締結記念 講演会 「ドイツと日本における学術情報流通基盤の未来」（於：学術総合センター　東京ドイツ文化センターとの共催）	
平成 23 （2011）	10/06　第 1 回運営委員会	10/28　第 1 回 SPARC Japan セミナー2011 Open Access Week 「OA 出版の現況と戦略（ジャーナル出版の側から）」 12/06　第 2 回 SPARC Japan セミナー2011 「今時の文献管理ツール」ワークショップ 01/31　第 3 回 SPARC Japan セミナー2011「学術情報流通の新たな展開 ― 研究者・学会とオープンアクセス ―」 02/10　第 4 回 SPARC Japan セミナー2011「学術情報流通の未来を切り開く ― 電子ジャーナルの危機とオープンアクセス ―」 02/29　第 5 回 SPARC Japan セミナー2011「OA メガジャーナルの興隆」 03/26　第 6 回 SPARC Japan セミナー2011「数学出版に関するワ	08/28-09/01 American Chemical Society (ACS) Fall 2011 National Meeting & Exposition（第 242 回米国化学会秋季大会）に出展(於：デンバー) 09/04-09 14th Asian Chemical Congress 2011 (14 ACC)（第 14 回アジア化学会議）に出展(於：バンコク) 10/26　2011 Open Access Korea(OAK) Conference での発表 (於：ソウル)

	03/27　第 2 回運営委員会	ークショップ」（於：東京理科大学　Project Euclid 主催、日本数学会共催ワークショップ）	
平成 24 （2012）	12/10　第 1 回運営委員会 03/26　第 2 回運営委員会	05/25　第 1 回 SPARC Japan セミナー2012「学術評価を考える」 06/19　第 2 回 SPARC Japan セミナー2012「ジャーナルの発展をもとめて～プラットフォーム移築を中心に～」 07/25　第 3 回 SPARC Japan セミナー2012「平成 25 年度 科学研究費補助金（研究成果公開促進費）改革」 08/23　第 4 回 SPARC Japan セミナー2012「研究助成機関が刊行するオープンアクセス誌」 10/26　第 5 回 SPARC Japan セミナー2012「Open Access Week ー 日本におけるオープンアクセス、この 10 年これからの 10 年」 12/04　第 6 回 SPARC Japan セミナー2012「オープンアクセスによって図書館業務はどう変わるのか～図書館のためのオープンアクセス講座～」 02/19　第 7 回 SPARC Japan セミナー2012「図書館によるオープンアクセス財政支援」	07/02-07 European Congress of Mathematics (ECM) に出展（於：クラクフ、ポーランド） 08/19-21 American Chemical Society (ACS) Fall 2012 National Meeting & Exposition（第 244 回米国化学会秋季大会）に出展　（於：フィラデルフィア） 08/26-30 4th EuCheMS Chemistry Congress（第 4 回ヨーロッパ化学会議）　に出展（於：プラハ） 12/26-27 RIMS 共同研究　研究会にて基調講演（於：京都大学）
平成 25 (2013)		06/07　第 1 回 SPARC Japan セミナー2013「SPARC と SPARC Japan のこれから」 08/23　第 2 回 SPARC Japan セミナー2013「人社系オープンアクセスの現在」 10/25　第 3 回 SPARC Japan セミナー2013「オープンアクセス時代の研究成果のインパクトを再定義する：再利用と Altmetrics の現在」 12/19　第 4 回 SPARC Japan セミナー2013「今日の問題を解く、学術情報の受信と発信－Think Globally, Act Locally」 02/07　第 5 回 SPARC Japan セミナー2013「アジアを吹き抜けるオープンアクセスの風―過去、現在、未来」	08/06　第 1 回　OA ジャーナルへの投稿に関する調査ワーキンググループ開催 10/02　第 2 回　OA ジャーナルへの投稿に関する調査ワーキンググループ開催 12/04　SCOAP[3] と MOU を締結 01/27　RIMS 共同研究　研究会にて基調講演（於：京都大学） 03/02　COAPI Meeting へ参加　（於：カンザスシティー）

	03/24　第 1 回運営委員会		03/03-04 SPARC2014 Open Access Meeting 本会議への参加 (於：カンザスシティー) 03/13　第 3 回　OA ジャーナルへの投稿に関する調査ワーキンググループ開催
平成 26 (2014)	01/15　第 1 回運営委員会 03/19　第 2 回運営委員会	08/04　第 1 回 SPARC Japan セミナー2014「大学/研究機関はどのようにオープンアクセス費用と向き合うべきか―APC をめぐる国内外の動向から考える」 09/26　第 2 回 SPARC Japan セミナー2014「大学における OA ポリシー：日本版 OA ポリシーのモデル構築に向けて」 10/21　第 3 回 SPARC Japan セミナー2014「「オープン世代」の Science」 03/09　第 4 回 SPARC Japan セミナー2014「グリーンコンテンツの拡大のために我々はなにをすべきか？」	05/21-23 COAR(Confederation of Open Access Repository) 2014 Annual meeting への参加　(於：アテネ) 06/09-13 OR2014 (The 9th Annual International Conference on Open Repositories)への参加　(於：ヘルシンキ)
平成 27 (2015)	08/04　第 1 回運営委員会 11/30　第 2 回運営委員会 03/24　第 3 回運営委員会	09/30　第 1 回 SPARC Japan セミナー2015「学術情報のあり方 - 人社系の研究評価を中心に - 」 10/21　第 2 回 SPARC Japan セミナー2015「科学的研究プロセスと研究環境の新たなパラダイムに向けて - e-サイエンス，研究データ共有，そして研究データ基盤 -」 01/19　第 3 回 SPARC Japan セミナー2015「研究者向けソーシャルメディアサービスの可能性」 03/09　第 4 回 SPARC Japan セミナー2015「研究振興の文脈における大学図書館の機能」	04/15-16 COAR(Confederation of Open Access Repository) 2015 Annual meeting への参加　(於：ポルト) 05/18-20 ORCID-CASRAI Joint Outreach Conference & Codefest 及び ORCID Board Meeting への参加　(於：バルセロナ) 06/08-11 OR2015 (The 10th Annual International Conference on Open Repositories)への参加　(於：インディアナポリス) 11/03-06 ORCID Outreach Meeting & Codefest, November 2015 及び ORCID Board Meeting への参加(於：サンフランシスコ) 02/02-03 ORCID Board Meeting への参加　(於：ロンドン) 03/07-08 SPARC Meeting on Openness in Research & Education への参加　(於：サンアントニオ)

平成 28 (2016)	09/06　第 1 回運営委員会 03/14　第 2 回運営委員会	09/09　第 1 回 SPARC Japan セミナー2016「オープンアクセスへの道」 10/26　第 2 回 SPARC Japan セミナー2016「研究データオープン化推進に向けて：インセンティブとデータマネジメント」 02/14　第 3 回 SPARC Japan セミナー2016「科学的知識創成の新たな標準基盤へ向けて：オープンサイエンス再考」	04/11-13 COAR (Confederation of Open Access Repository) Annual Meeting 2016 への参加　(於：ウィーン) 05/18-19 ORCID Board Meeting への参加　(於：トロント) 06/08-11 CRIS2016 (The 13th International Conference on Current Research Information Systems)への参加 (於：セントアンドリュース) 06/13-16 OR2016 (The 11th Annual International Conference on Open Repositories)への参加　(於：ダブリン) 09/12-13 arXiv.org Member Advisory Board Meeting への参加 (於：イタカ) 09/15-17 RDA (Research Data Alliance) 8th Plenary Meeting への参加　(於：デンバー) 10/14-15 SCOAP³ Executive Meeting への参加(於：ジュネーブ) 11/09-10 PIDapalooza (PID：Persistent ID) への参加 (於：レイキャビック) 12/07-09 SCOAP³ に関する講演・意見交換会への協力 (於：NII、京都) 01/17-19 IDF Strategic Meeting への参加　(於：バルセロナ) 01/25-27 Open Access に係る海外機関の調査 (於：レーゲンスブルク，ミュンヘン) 03/20 SCOAP³ Executive Meeting への参加(於：ジュネーブ) 03/23-24 SCOAP³ Governing Council Meeting への参加 (於：ジュネーブ)
平成 29 (2017)	09/26　第 1 回運営委員会	09/13　第 1 回 SPARC Japan セミナー2017「図書館員と研究者の新たな関係：研究データの管理と流通から考える」 10/30　第 2 回 SPARC Japan セミナー2017「プレプリントとオープンアクセス」	04/05-07 RDA (Research Data Alliance) 9th Plenary Meeting への参加　(於：バルセロナ) 05/08-10 COAR (Confederation of Open Access Repository) Annual meeting 2017 への参加　(於：ヴェニス) 06/13-14 IDF Strategic Meeting International DOI Foundation への参加　(於：テジョン) 06/15 DOI Outreach Meeting への参加　(於：ソウル) 06/27-30 OR2017 (The 12th Annual International Conference on Open Repositories)への参加　(於：ブリスベン) 09/18-21 RDA (Research Data Alliance) 10th Plenary Meeting への参加　(於：モントリオール) 09/26 SCOAP³ Governing Council Meeting への参加(テレビ会議) 10/05 arXiv.org Member Advisory Board Meeting への参加　(於：ニューヨーク) 12/07 SCOAP³ FORUM (Webinar)

	03/19　第 2 回運営委員会	02/21　第 3 回 SPARC Japan セミナー2017「オープンサイエンスを超えて」	01/23-24 PIDapalooza (PID：Persistent ID) への参加 (於：ジローナ) 01/25-26 IDF Annual Meeting への参加　(於：バルセロナ)

５ 刊行物一覧

５．１ 国際学術情報流通基盤整備事業(SPARC Japan)年報

〔日本語〕

・国際学術情報流通基盤整備事業（SPARC Japan）年報 平成 28（2016）年度
http://www.nii.ac.jp/sparc/publications/annual/pdf/sparc_annual_2016.pdf

〔英語〕

・SPARC Japan (International Scholarly Communication Initiative) Annual Report FY2015
http://www.nii.ac.jp/sparc/en/publications/pdf/sparc_annual_2015-E.pdf

５．２ SPARC Japan ニュースレター

〔日本語〕

・SPARC Japan NewsLetter 第 33 号（2017 年 12 月）
http://www.nii.ac.jp/sparc/publications/newsletter/PDF/sj-NewsLetter-33.pdf

・SPARC Japan NewsLetter 第 34 号（2018 年 3 月）
http://www.nii.ac.jp/sparc/publications/newsletter/PDF/sj-NewsLetter-34.pdf

・SPARC Japan NewsLetter 第 35 号（2018 年 7 月）
http://www.nii.ac.jp/sparc/publications/newsletter/PDF/sj-NewsLetter-35.pdf

〔英語〕

・SPARC Japan NewsLetter No. 32, Apr. 2017
http://www.nii.ac.jp/sparc/en/publications/pdf/sj-NewsLetter32E.pdf

・SPARC Japan NewsLetter No. 33, Dec. 2017
http://www.nii.ac.jp/sparc/en/publications/pdf/sj-NewsLetter33E.pdf

５．３ SPARC Japan セミナードキュメント

【第 1 回 SPARC Japan セミナー 2017】（平成 29 年 9 月 13 日）

「図書館員と研究者の新たな関係：研究データの管理と流通から考える」
https://www.nii.ac.jp/sparc/event/2017/20170913.html

・「開会挨拶/趣旨説明」 能勢　正仁 （京都大学大学院理学研究科）

・「研究者にとってのデータの意味と大学におけるデータ管理への期待」
倉田　敬子 （慶應義塾大学文学部）

・「学術リポジトリは研究者と図書館員を繋げるのか？」
大澤　剛士 （農業・食品産業技術総合研究機構農業環境変動研究センター）

・「研究データ管理の組織的支援と図書館の役割について」
西薗　由依 （鹿児島大学/JPCOAR 研究データタスクフォース）

・「新たな学術情報流通において JPCOAR スキーマが果たす役割」
片岡　朋子 （お茶の水女子大学/JPCOAR メタデータ普及タスクフォース）

・「研究データ利活用協議会(RDUF)紹介」

武田　英明　（国立情報学研究所/研究データ利活用協議会）

・「全体議論」

〔モデレーター〕能勢 正仁　（京都大学大学院理学研究科）

【第 2 回 SPARC Japan セミナー 2017】（平成 29 年 10 月 30 日）

「プレプリントとオープンアクセス」

https://www.nii.ac.jp/sparc/event/2017/20171030.html

・「開会挨拶/趣旨説明」

坊農　秀雅　（情報・システム研究機構ライフサイエンス統合データベースセンター）

・「arXiv.org の次世代システムの公開と戦略」

引原　隆士　（京都大学図書館機構長/arXiv.org 会員コンソーシアム代表）

・「学術情報共有とオープンアクセスの未来」

Gregg Gordon　（Managing Director, SSRN）

・「化学分野におけるプレプリントの位置付け・課題等について」

生長　幸之助　（東京大学大学院薬学系研究科/化学ポータルサイト Chem-Station 副代表）

・「生命科学分野におけるプレプリントの位置付けや経験について，統合 TV について」

小野　浩雅　（情報・システム研究機構ライフサイエンス統合データベースセンター）

・「全体議論」

〔モデレーター〕坊農 秀雅　（情報・システム研究機構ライフサイエンス統合データベースセンター）

【第 3 回 SPARC Japan セミナー 2017】（平成 30 年 2 月 21 日）

「オープンサイエンスを超えて」

https://www.nii.ac.jp/sparc/event/2017/20180221.html

・「趣旨説明」 蔵川　圭　（国立情報学研究所）

・「オープンサイエンスを真に理解する：その有益性の潜在能力，脆弱性，機能的パフォーマンスの問題，これらの解決策を講じない方法」

Paul A. David　（Stanford University）

・「データ駆動型の科学研究エコシステムとしてのオープンサイエンス
ー過去のコミュニティ実践事例と日本の視点」

村山　泰啓（情報通信研究機構/ICSU-World Data System）

・「オープンリサーチを可能にするには」

Heather Joseph　（SPARC North America）

・「デジタル時代の研究プロセスと大学，大学図書館における支援のあり方」

倉田　敬子　（慶應義塾大学文学部）

・「デジタル化時代の研究者のために図書館が構築すべき学術情報環境」

市古　みどり　（慶應義塾大学三田メディアセンター）

・「全体議論」

モデレーターより：「討論への問題提起：オープンサイエンスの知識論を考える」

〔モデレーター〕深貝　保則（横浜国立大学大学院　国際社会科学研究院）

SPARC Japan NewsLetter
ISSN 1883-826X
NO.33 2017 年 12 月

SPARC Japan ニュースレターでは、各回セミナーの報告に講演やパネルディスカッションを書き起こしたドキュメントを加え、さらにそのほかの SPARC Japan の活動をご紹介します。

また、6 月 1 日付で SPARC Japan のロゴを SPARC(US)とあわせて変更したため、本ニュースレターも今号から新しいロゴにしました。旧ロゴの新しいアイディアの誕生や知識のオープンシェアリングを拡大するといった”spark”のイメージを引き継ぎつつ、シンプルになりました。アスタリスクに似たシンボルは、デジタル環境における SPARC の確固とした立場を表現し、そのエネルギーとダイナミズムを強調しています。

SPARC Japan ➡ SPARC Japan

CONTENTS

■ SPARC Japan 活動報告
SPARC Japan 運営委員会
海外動向調査

■ SPARC Japan セミナー報告
概要
参加者から
企画後記
ドキュメント
（講演・全体議論）

■ SPARC Japan 活動報告

CC BY

SPARC Japan 運営委員会

SPARC Japan 運営委員会の会議資料をウェブサイトで公開しています。
http://www.nii.ac.jp/sparc/about/committee/

海外動向調査

第 5 期 SPARC 事業の達成目標の一つとして「オープンアクセスに関する基礎的情報の把握」を掲げており、その具体的活動として海外動向調査を行っています。今回は国際会議の参加報告をご紹介します。

- COAR 2017 年（第 8 回）総会参加報告
  お茶の水女子大学　香川 朋子、国立情報学研究所　山地 一禎（以上、JPCOAR メタデータ普及タスクフォース）
  2017 年 5 月 8~10 日
  http://id.nii.ac.jp/1458/00000036/

- Open Repositories 2017 参加報告
  九州大学　林 豊（JPCOAR メタデータ普及タスクフォース）
  2017 年 6 月 27~30 日
  http://id.nii.ac.jp/1458/00000030/

- 14th International Conference on Digital Preservation (iPRES2017) 参加報告
  国立極地研究所　南山 泰之（JPCOAR 研究データタスクフォース）
  2017 年 9 月 26~27 日
  http://id.nii.ac.jp/1458/00000033/

■ SPARC Japan セミナー報告

CC BY

第 1 回 SPARC Japan セミナー 2017

「図書館員と研究者の新たな関係：研究データの管理と流通から考える」

2017 年 9 月 13 日（水）　国立情報学研究所 19F 会議室　参加者：60 名

本セミナーでは、図書館員と研究者が研究データをいかに管理・流通させていくかという視点を共有し、その新たな関係構築について考えるため、研究者から現状分析を踏まえた図書館が関与する可能性について、図書館員から RDM トレーニングツールによる研究者支援等について紹介しました。

次ページ以降に、当日参加者のコメント（抜粋）、企画後記およびドキュメント全文（再掲）を掲載しています。その他の情報は SPARC Japan の Web サイトをご覧ください。(http://www.nii.ac.jp/sparc/event/2017/20170913.html)

CC BY

概要

日本における機関リポジトリは、公開されているものだけでその数 700 を超えており、世界でも類を見ないくらい大きな規模に成長してきた。こうした機関リポジトリは、主に大学の教育研究活動の成果である学術情報の収集・組織化・保存及び発信といった役割を担い、オープンアクセス推進の一翼を担ってきたが、更なる推進にあたっては、研究活動の過程で生み出される研究データについても同様の仕組みを構築する必要がある。

折しも平成 28 年 7 月には、日本における機関リポジトリを振興・相互支援することを目的として、「オープンアクセスリポジトリ推進協会（Japan Consortium for Open Access Repository : JPCOAR）」が設立され、研究データを含む学術情報の管理・流通の発展が期待されるところである。オープンアクセスリポジトリを効果的に運用していくためには、図書館員にとってはこれまであまり触れる機会がなかった研究データを管理する方法やリポジトリを通した図書館による新たな研究者支援手段を理解する必要がある。また、研究者にとっては、研究データのオープン化が生み出す新たな価値を認識し、図書館員との連携によりリポジトリの積極的・自発的な活用を目指していく必要がある。

以上の状況を踏まえて本セミナーでは、図書館員と研究者が、研究データをいかに管理・流通させていくかという視点を共有し、そこからもたらされる両者の新たな関係について考えてみたい。

参加者から

（大学/図書館関係）
・すぐに解決策が見つかる問題ではありませんが、研究者の方の率直なご意見を伺えたのが非常に興味深く、参考になりました。逆に、研究者の方も図書館員の生の声を聴けて良かった、とおっしゃっていたので、研究者と図書館員をつなぐこのような企画はとても意義があると思いました。

（大学/研究者）
・図書館員の職種と研究者の職種によるカルチャーの違いを感じました。

（その他/図書館関係）
・Embedded Librarian や学術情報支援部門としての役割など研究図書館の活かし方、未来像を感じさせる有意義な講演会でした。ただ、図書館員から、研究者からお仕事を依頼されること自体が至上の喜びで、共同研究者にはなれない、外部資金はいらない、自分のした仕事への評価もいらない等、驚きのコメントがされていたことに残念な思いがしました。私は研究者のプロジェクトの中の業務を分担されて行なったこともありますし、資料収集も依頼されることがあります。そうしたときに謝辞や協力者に名を載せて頂くことがありますが、自分のした仕事が役に立った評価を受けたということを素直に嬉しく感じますし次の仕事へのモチベーションにもなります。また当館の二代前の館長は当館所蔵の古文書を研究材料とする研究者と共同研究者として名を連ねたこともあります。完璧な Embedded Librarian には程遠いですが、図書館員の考え方しだいで研究者との距離は縮まる、お互いの仕事に上下を付けずに尊重しあって対等に対話できるようになるのではないかと思います。

JPCOAR に携わっている方々には大変感謝致します。スキーマの策定、ありがとうございます。研究者と図書館員の間の風通しが良くなることもオープンサイエンスの一部なのかもと目から鱗の講演会でした。

（その他/研究者）
・特に大澤先生の議論のディスカッションが役立った。職分によってのインセンティブの違いについては実行面で重要なのではっきりさせてもらえてよかった。

企画後記

研究データを公開するうえで必要な手続きについて、図書館員と研究者の考え方やインセンティブが異なることがよくわかりました。図書館員の方の生の声が聞けたのは、質疑応答の時間を普段より長くとるように企画をした者にとって大きな収穫でした。発言くださった図書館員の方に深く感謝申し上げます。図書館員と研究者のギブアンドテイクがうまく進んでいくことを願っています。

能勢　正仁
（京都大学大学院理学研究科）

今回のセミナーの企画並びに当日の twitter 係を担当いたしました。研究データの管理と流通について図書館が組織として関わる際、その貢献はいかにして表されるべきなのか、研究者と図書館員のとらえ方の差異が浮き彫りになったのではと考えています。ユーザである研究者から共著や謝辞にて貢献を表される機会が増えることが、図書館員のさらなるインセンティブに、そして組織の評価の向上にもつながることを期待しています。

林　賢紀
（国立研究開発法人 国際農林水産業研究センター）

図書館員としては、研究データ管理にかかる新たな役割の示唆、あるいは研究者の立場からの率直な疑問・意見など、得るものが多かった一方で、自身のインセンティブは何か、それを享受するに足るものを提示できるか、ということもより深く考える機会になったのではと思います。RDM トレーニングツールやメタデータスキーマなどの活用を基盤に、上記のような、あるいはさらに別の協働の姿へ進む、そのための足掛かりとなれば幸いです。

中谷　昇
（鳥取大学附属図書館）

第 1 回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係:研究データの管理と流通から考える」

開会挨拶/趣旨説明

能勢 正仁

（京都大学大学院理学研究科）

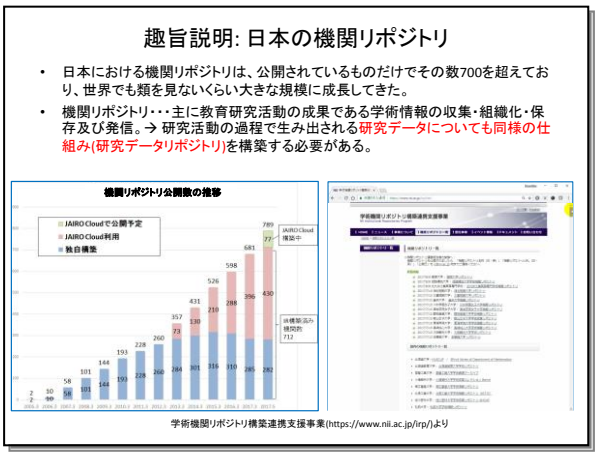
能勢 正仁

1998年に京都大学理学研究科で博士（理学）取得後、米国ジョンズホプキンス大学でポストドクトラルフェローとして3年間研究を行う。2001年帰国、現職。専門は、超高層物理学、地球電磁気学。主な研究テーマは、地磁気変動・脈動、内部磁気圏の高エネルギー粒子ダイナミクス、サブストーム、地磁気指数など。最近は、科学データへデジタルオブジェクト識別子を付与する活動にも積極的に関わっている。

日本の機関リポジトリは、公開されているものだけでその数は 700 を超えており、世界でも類を見ないくらい大きな規模に成長しています。

図 1 のグラフは、機関リポジトリの公開数の推移です。2017 年 5 月時点で 789 となっており、たくさんのリポジトリが日本では公開されています。

機関リポジトリは、現在では、教育研究活動の成果である学術情報の収集・組織化・保存および発信という役割を担っています。学術成果である論文や研究発表スライドなど、主に文書を保管しています。



（図 1）

研究データリポジトリの必要性

次のステップとして、研究活動の過程で生み出される研究データについても、同様の仕組みを構築する必要があると考えられます。これを「研究データリポジトリ」と呼びます。日本は、約 1,200 ある大学のうち、約 7 割がリポジトリを持っているので、次のステップに進む上で非常に有利だと私は考えています。

とはいえ、研究データリポジトリにすぐ推移できるかというと、まだ壁があります。研究データリポジトリを取り巻く図書館員と研究者の現状を私が理解できる範囲でまとめたものが図 2 です。図書館員は、既に機関リポジトリの運営の経験がある、機関リポジトリ推進委員会の活動を通したノウハウを持っている、オープンアクセスリポジトリ推進協会（JPCOAR）が新たに始まったなどの理由から、研究データリポジトリに対する準備は進んでいます。

一方、研究者は、研究データリポジトリにネガティブなイメージを持っています。研究公正のために研究データをきちんと管理するようにという上からの押し付けがあったり、これはまだ日本では一般的ではあり

ませんが、欧米では既に、研究計画書と合わせて、得られたデータの管理方法を定めたデータ・マネジメント・プラン（DMP）を提出することが一般的になっており、さらに書類が増えるというイメージがあったりします。

研究データリポジトリといっても、これが実際に稼働している大学は現実にまだないので、そうすると研究者は、自前のデータベースを持つ、自前のサーバー上でデータを管理するという手続きになってきます。かつ、そのためにメタデータをつくり、それを管理しないといけないという仕事が増え、結局、研究、教育、雑務に忙しくて、データのオープン化まで手が回らないという状況にあると考えられます。

研究データリポジトリができて、図書館員の側で研究データを管理することが可能になったとしても、まず研究データとは一体どういうものか、どう扱うのかというのが共通の疑問点であると思われます。

また、図書館員とお話しすると、共通して、「研究者のニーズが本当にそんなにあるのか」「大学図書館に関わり得る余地があるのか」という疑問が呈されます。

研究者と図書館との関わりは論文や図書がやはり中心で、一般的な研究者の認識として、データをオープンにする際に図書館員との連携が頭に浮かばないという問題や、データをオープンにするインセンティブがないという問題もあります。

「インセンティブがない」、私はこれを一つキーワードと考えています。インセンティブがないので、データリポジトリが進んでいくのにはいろいろな障壁があると思われます。こういったことから、図書館員と研究者の間で、それぞれどういうニーズがあって、どういう協力関係ができるかということの対話の場が必要だと考えました。以上のような経緯で今回の企画をさせていただきました。

図書館員と研究者との対話

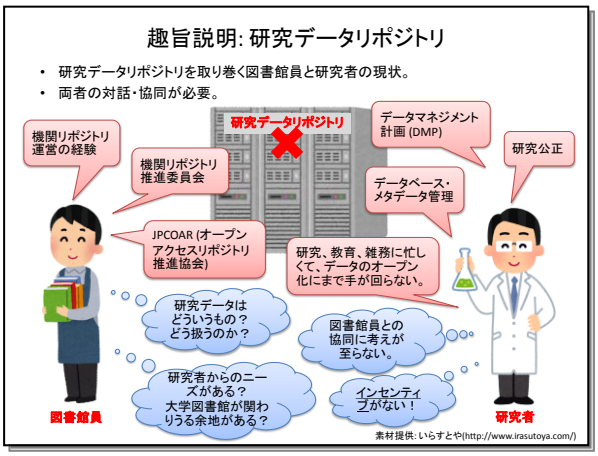
今回は、図書館員と研究者の置かれている状況について多くのことをご存じの登壇者の先生方にお話をしていただきます。

倉田先生には、研究者が抱えている現状について、広い範囲の話をしていただきます。

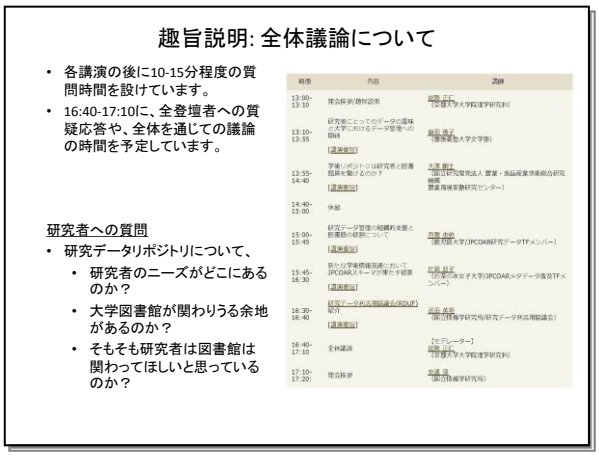
大澤先生には、研究者から図書館員に対して、どういうアプローチ、問い掛けができるかというお話をしていただけると思います。

休憩を挟んだ後は、特に JPCOAR に焦点を当て、西薗さんと片岡さんから JPCOAR の活動について、現在、図書館員の間でどういう話し合いが進んでいるか、ご報告を頂きたいと考えています。

両者の対話・協働が必要ということで、最後にいつものパネルディスカッションではなく、全体議論の時間を 30 分用意しています。前半の約 15 分は登壇者への質問、後半の約 15 分は特に研究者への質問という形でお願いしようと思っています。



（図 2）



（図 3）

図書館員は、研究者はリポジトリに対してどのようなニーズがあるのか、大学図書館が関わり得る余地があるのか、そもそも図書館に関わってほしいと思っているのかということを共通の疑問としていつもお持ちです（図 3)。ですから、最後の時間に研究者の方々から自発的に手を挙げていただいて、図書館員の方がたくさんいらっしゃるこの場で、ご自身のお考えを発言していただけると幸いです。

第 1 回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係:研究データの管理と流通から考える」

研究者にとってのデータの意味と大学におけるデータ管理への期待

倉田 敬子
（慶應義塾大学文学部）

講演要旨

研究データに対して研究者の持つ意識や行動は多様である。例えば、自分のデータだ、データだけ公開されても使えない、専門の研究者ならデータを見れば理解できる、など単純にデータ共有に積極的と消極的という二分法では理解できない。それは研究データが研究サイクルに埋め込まれたものであり、論文のように研究成果としての最終形が確定するわけではないからである。本発表では、研究者の意識の多様性とデータの特性について紹介した上で、大学でのデータの管理、保管に関する体制整備が進んでいない状況を報告し、大学等が目指すべき方向性を検討する。

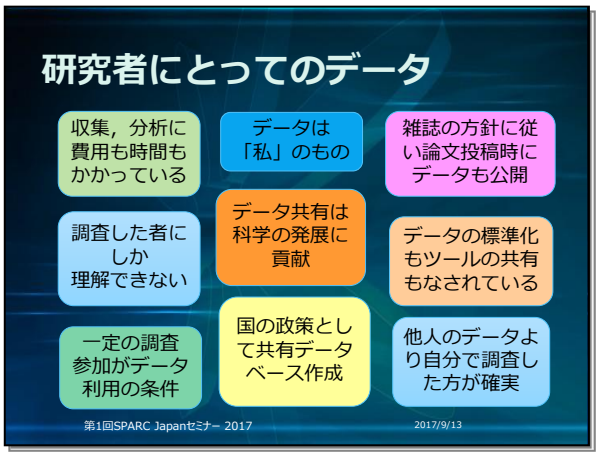
倉田 敬子

1987年慶應義塾大学文学研究科博士課程修了。1988年慶應義塾大学文学部図書館・情報学科助手、1993年同助教授、2001年から文学部図書館・情報学専攻教授。専門は学術コミュニケーション、特に情報メディアのデジタル化、オープンアクセス、研究データに関心を持っている。著書に『学術情報流通とオープンアクセス』（勁草書房）など。2008年度日本図書館情報学会賞授賞、2010年科学技術への顕著な貢献2010（ナイスステップな研究者）選定。

私からは、研究者がデータというものに関してどう思っているのか、それに対して図書館側が何をできるのか、それに対してどのような期待ができるのかをお話しします。

1.データ共有の阻害要因と推進要因

研究者がデータの共有や公開、もしくは研究データそのものに対して持っている意識は大変多様です（図 1）。もちろんデータ共有は科学の発展に貢献すると主張し、自らも実践している方もいますが、一方で研究者には「そう言われてもデータは私のものだよね」という感覚がやはりどこかにあります。自分が関わり、収集や分析に多大な費用も労力もかけたデータ、それを他人に見せることに、何となく抵抗があるのです。ただ、これは分野によって全く異なります。例えばヒトゲノムを中心とする生物医学分野では、「そんなものは出



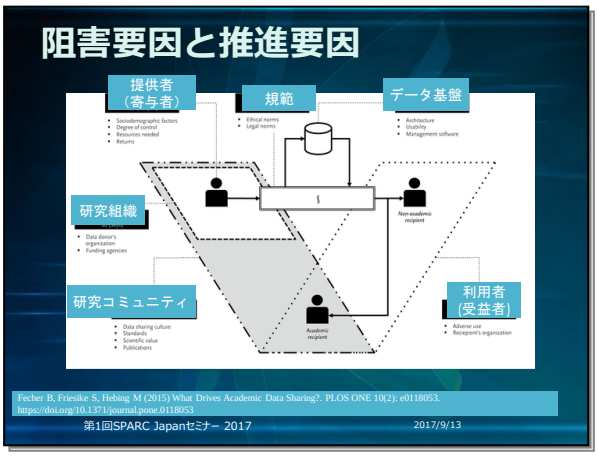
（図 1）

すのが当たり前でしょう」という感覚があり、一方他分野では「いやいや、そうは言ってもね」という感覚が交ざっている段階だと思います。

データ公開に対する意識や行動の多様性が一体どういう違いから出てくるのかは、みんなが不思議に思うところです。これは研究者個人の意識の問題なのか、分野の違いなのか、投稿雑誌や助成機関が方針を義務化すれば一気に進むものなのか、年齢や業績などが関係するのか。阻害要因と推進要因について、幾つかの研究がなされています。これはあくまでも一例として見ていただきたいのですが、Fecher らのグループが、2,000～3,000 本あるデータ共有に関する論文の中から 98 本を選んでシステマティックレビューを行っています。さらに、ドイツの SOEP という、社会調査関係のデータを第三者の利用を前提として公開しているところのパネルデータを、実際に利用している人たちがどう思っているかを聞いた簡単なアンケート調査があります。Fecher らは、この二つを分析し、阻害要因と推進要因を図 2 のようにまとめています。

データ共有を考えるに当たっては、データを提供する、データの公開に寄与する者と、そのデータの受益者・利用者がいます。これは当然のことだと思います。ただし、この 2 者は別のグループにいるわけではなく、同じ分野もしくは違う分野であるかもしれませんが、基本的には研究者という同じグループにいて、どちらの立場にもなり得るのです。

彼らの行動をある意味では制約していて、推進・阻害の両方の面で効いてくるのが、研究組織がどういう方針を持っているか、研究コミュニティ全体がどういう意識であるかであり、それらに通底するのが規範です。研究倫理だけではなく、発表や公開をどう考えるかという一般的な規範があるということを示している図だと私は解釈しています。



（図 2）

私は、「データ基盤」というところが新しい要素になってくるのではないかと考えます。どのような基盤があるかということが、データ公開の促進や阻害に影響してくる。言ってみれば、何も基盤がないところでやれと言われてもなかなか進まないわけで、ここに私は一番関心を持っています。

2.研究者の意識・行動のモデル化

このように阻害要因と推進要因の研究は随分進んでいるのですが、ここからは私の研究の一部を紹介します。私どもの研究グループは、研究データやオープンサイエンスについて、研究者がどのように行動し、どのような意識を持っているのかに大変興味を持っています。

データの共有や公開に関して、日本人研究者がどう考えているのかを知るために、自然科学系の研究者 23 人を対象にインタビュー調査を行いました。ただ、対象者は年齢も業績も上の方が多くなってしまったので、若い方の意見が十分でない可能性はあります。研究活動のどのようなプロセスの中でデータを扱っているのか、データ共有をどう考えるのかというインタビューを行い、その内容を分析しました。

しかし、内容分析の結果は、本当にいろいろな意見があり、まとまりがなくなってしまったので、特に研究データに関して言及している言説だけを 406 件抽出し、クラスター分析をかけました。そうすると 14 クラスターに分けることができました。

ここからは私達の方で解釈したものですが、それを五つのグループに分けられるのではないかと考えました。図 3 の右側が論文に出した図です。「Research Data」が真ん中にあり、それを囲っているイエローと

ピンクとグリーンの部分が、研究者がリサーチデータに対して取る、最も表面に表れる三つの態度だと考えました。1）データを共有しようと思う、2）共有したくないと思う、3）単に積極的・消極的には分類できない、データをどう考えるべきか、データにはどのような意味があるのか、研究にとってこのデータは何なのか、科学にとってデータとは何かという、意味付けの問題に関して語られたことです。そこのデータの意味付けの問題は、単にオープンやリストリクティブとは別の意識として出てくるのではないかということを、このモデルで示したかったということです。

その背景で、実際に研究活動においてデータをどう処理し、分析し、それを保管しているか、広範囲なさまざまな活動に関して語られました。研究者は基本的にはデータを保管していて、それをオープンな形にしていくかというところに関しては、さまざまな活動があるということだと思います。そして、先ほど申しましたように、私はその裏に、データの基盤に関わる部分があるのではないかと考えています。その分野においてデータリポジトリがあるということはもちろん直接的な基盤ですが、それだけではなく、国の政策や、よく投稿する雑誌など、そういうことも全部含めて、裏にあるインフラ的なものが言説の中からも出てきたように考えています。この図はある種の構造化といいますか、いろいろな観点が入り込んで、今の研究者の意識や行動というものを構築しているのではないかということをお示ししたかったのです。

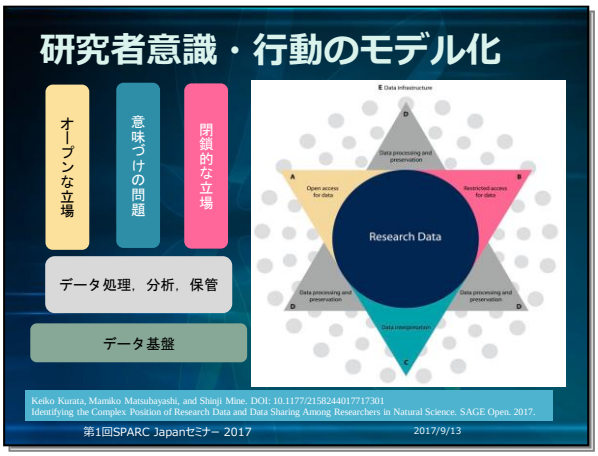
3.データとは何なのか？

研究者の意識・行動が複雑・多様になる大きな原因は、「データ」とは何なのかが分からないということです。厳密に言うと、研究者はデータを扱っているのですが、それが何かを第三者に説明するのが非常に難しいのです。

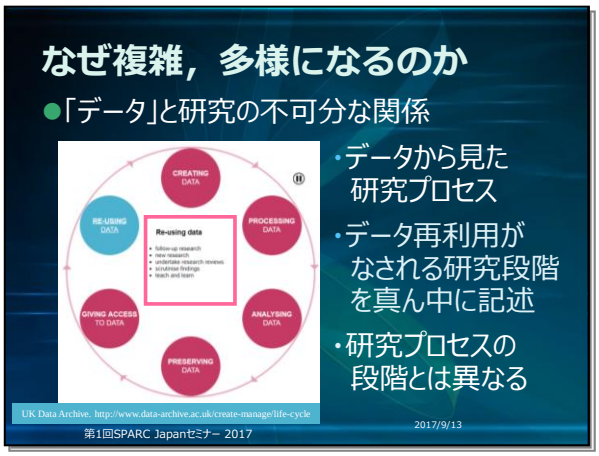
一般に考えると、調査データ、観察データ、実験データはデータだと思います。でも、それだけではないはずで、実験データに付随する形で、実験機器に関するデータ、実験した条件に関するデータも必要です。これをメタデータと言う場合もあります。では例えばそれを解析するために使ったソフトウエアはどうなるのか。会合記録はどうなのか。スライドはデータではないのか。さらに、アイデアのメモまでデータに含めるという考え方をする人もいます。一体どこをデータとして切り取るのかということ自体が大きな問題になるのです。

図 4 は日本の講演などでもよく利用されている、英国のデータアーカイブの図で、データのライフサイクルを表しています。データがつくられて、処理されて、分析されて、何らかの形で実際に表に出て、それに対するアクセスがあって、最後にもう一回使われるというサイクルで、データの再利用とは何かという細かい記述が真ん中にあります。

これは、データから見た研究プロセス、データと研究との関係を描いた図としては、大変よくまとまっていると思います。ただ、これは研究者から見た研究プ



（図 3）



（図 4）

ロセスとは異なるのです。

もう一つ例を出します。図 5 はリサーチライフサイクルとデータサイクルとを結び付けようと考えられた図で、よく考えられてはいるのですが、そのリサーチライフサイクルとしては、アイデアができて（ideas）、プロジェクトをつくって、人を集めて（partners）、申請して（proposal writing）、お金を取って、実際の研究をして（research process）、その成果を発表する（publication）というものになっています。確かにこれは研究のサイクルです。そしてこの図では、research process の中でデータがどのように動くかを見ています。しかし、やはりここだけでは不十分なのです。もっと前の段階でもデータを使うでしょうし、後の段階でのデータの公開はどうなるのかということがあります。

つまり何が言いたいかというと、研究プロセスを中心に考えると、データのライフサイクルがあまりうまく表示できないし、データのライフサイクルを考えると、研究とどう関わってくるのかが見えてこないということです。研究者が考えるデータの扱いと、データの管理を考える側のデータの扱いをどうやってうまく擦り合わせていくのかということが問題になるのではないかと考えています。

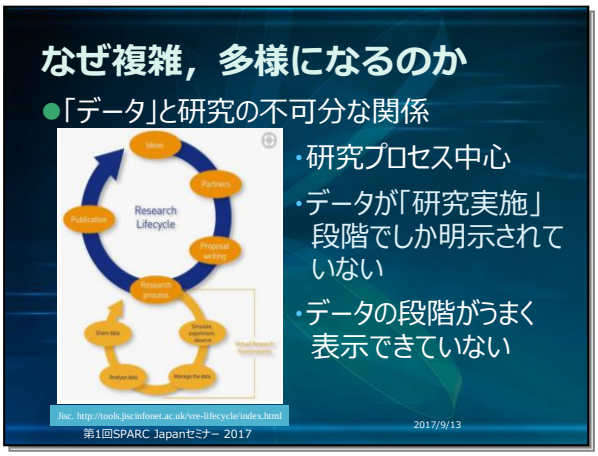
そう考えると、研究データを管理するには当然のことながら研究プロセスへの理解が必要です。研究者がどのような段階で何を求めていて、だから、今このデータなのだということを理解した上で、データの管理を考えなければいけません。データの管理は、研究支援全体の文脈の中で考えなければいけません。データだけを「このような形できちんとしてください」と言っても、それはやはりなかなか難しいわけです。なぜそのデータをそのような形で管理していかなくてはいけないのか、例えばデータ管理計画をなぜこんなときに出さなければいけないのかというのは、研究プロセス全体とデータの管理全体の両方をよく理解しないと、研究者と管理側の両方の考え、認識を一致させることはできないだろうということです。

4.大学・研究機関における支援の現状

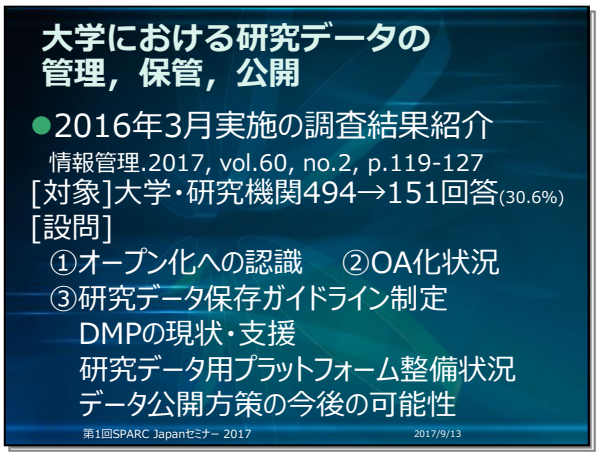
この研究支援全体を考えるということは、図書館だけでできることではないと思います。では今、大学は研究データについて何を行っているのかを知るために、大学における研究データの管理・保管・公開に関して、2016 年 3 月に質問調査を行いました（図 6）。この結果は「情報管理」に掲載されているので、詳細はそちらを見ていただければと思います。対象は、大学・研究機関 494 機関を選び、151 件の回答を得ました。

基本的な設問は、オープンアクセスおよびオープンサイエンスに関するさまざまな政策を知っているか、オープンアクセス方針をつくっているか、機関リポジトリを構築しているかなどです。

その上で中心となるのは、研究データ保存のためのガイドラインを既に制定しているかという問いです。これは、「研究活動における不正行為への対応等に関



（図 5）



（図 6）

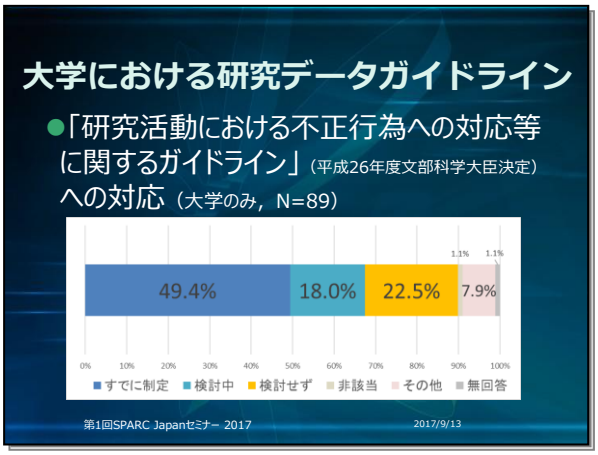
するガイドライン」を制定しなければいけないという平成 26 年度文部科学大臣決定があり、それに従っているかどうかということです。それから、データ・マネジメント・プラン（DMP）は現状どうなっているか、支援がどうなっているかということを聞いています。

幾つか結果をご紹介します。「研究活動における不正行為への対応等に関するガイドライン」への対応については、大学が対応するものなので、大学のみのデータを出しています（図 7）。

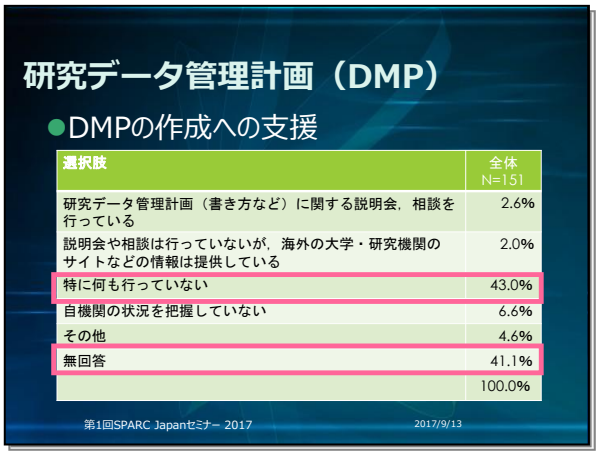
「すでに制定」が 49.4％と約半分、「検討中」が 18.0％で、検討中も含めると、7 割が基本的にはガイドライン自体は制定しているということです。ただし、慶應義塾大学をはじめとする幾つかのガイドラインを見ると、基本的には研究者に丸投げなのです。保管はしろ、けれども大学は特に何もしない、それでも一応ガイドラインはつくったということになります。最低限の認識はあるし、研究不正に対処する形での対応はしている、でも DMP については全く知らず、DMP の提出を義務付けているという回答は一つもありませんでした（図 8）。DMP の作成を推奨しているところもごくわずかです。

一番問題なのは、DMP 自体を知らないと回答しているところが、特に研究機関ではほぼ半分であることです。自分の機関の DMP の状況がどうなっているかを把握していないというのも問題だと思います。この二つが回答で全体のほとんどを占めていました。

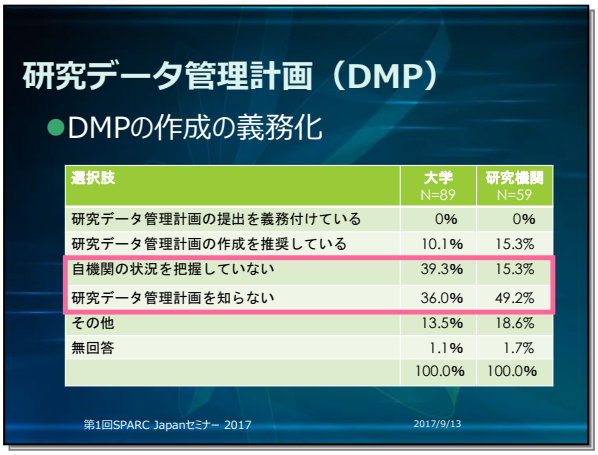
当然のことながら、DMP 作成支援は特に何も行っていないという回答が 43.0%、無回答が 41.1%でした（図 9）。無回答というのは、分からないから答えようがなかったということだと思います。大学は今のところほとんど何もやっていないと言えます。



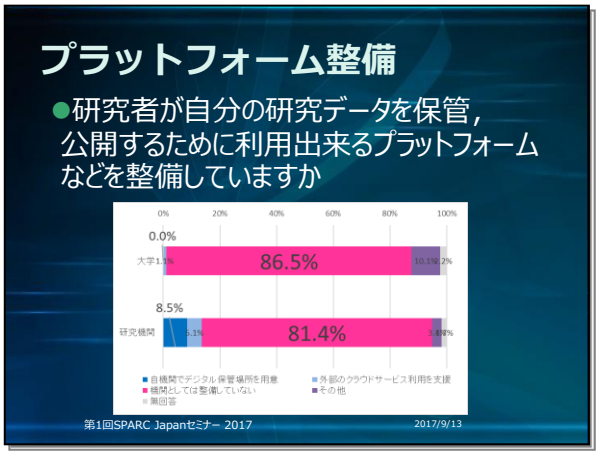
（図 7）



（図 9）



（図 8）



（図 10）

また、「研究者が自分の研究データを保管、公開するために利用できるプラットフォームなどを整備していますか」という質問に対しては、図 10 の帯グラフの上が大学、下が研究機関ですが、どちらも 8〜9 割が何もやっていないということでした。

これらのことから、物理的にも制度的にも支援ができていないというのが、残念ながら現状だと考えられます。

また、この調査は非常に低い回収率でした。その大きな原因は、大学宛てに調査票を出してしまったことだと考えられます。回答する担当部署がなく、誰が答えていいかも分からなかったのでしょう。調査票が図書館に行って、図書館に非常に意識の高い方がいればきちんと回答してくださったかもしれませんが、どこかに行ってしまったということも含めて、回収率が低かったのだと思います。

結論としては、大学・研究機関はオープンデータに関して基本的には全く知識がなく、データの保管はあくまでも「研究不正」の文脈としてしか認識していない。保管場所の提供もなく、何らサポートもなされていないということが分かりました。

5.図書館に何ができるのか

この状況で、図書館だけ頑張れというのはちょっと酷な話だと思います。大学全体がそのような状況にある中では、図書館員にたとえ意識の高い方がいたとしても、オープンデータの取り扱いは伝統的な図書館業務と見なされることがなかなかなく、図書館の中で理解を得ることも難しいと思われます。そこが一つ大きな問題です。「こういうことをやってみたい」「こういうことが必要だ」と思う方が、図書館の中ですら認められない可能性があるのです。

ただ、今の大学図書館をめぐるさまざまな状況を考えると、この状況に対して何らかの役割を果たせると自ら主張していかないと厳しいのではないかと思います。

新しい役割の可能性としては、「①Embedded Librarian」と「②学術情報支援部門の構築」という二つの方向性があるのではないかと考えています。

5-1.Embedded Librarian

Embedded Librarian とは、日本ではほとんど定着しなかったサービスですが、欧米の医学図書館を中心とした研究図書館において一部ではありますが、研究・教育プロジェクトチームの一員となる図書館員のことです。図書館の中にいるのではなく、研究プロジェクトの会議その他に全て参加して、チームの一員として、基本的には文献探索と情報提供を行う専門職として認められています。先生方も忙しいので、「ちょっとこの辺の文献を探しといて」ということから始まって、さまざまな研究に関連する情報を探して提供します。医学分野などでは、システマティックレビューのようなものが必要なときには、Embedded Librarian が山のような文献をきれいに整理したり、要約を作成したりするなど活躍して、うまくいっている例はたくさんあります。

基本的にはそういう文献探索、情報提供の専門家のことを Embedded Librarian、Information Specialist のように呼んでいましたが、今これが学術コミュニケーション、データ管理を行える専門家として研究チームの一員となるという方向性が議論されています。

今、欧米では Librarian の職種が区分けされていて、「Scholarly Communication Librarian」という言葉が出はじめています。ジョブディスクリプションで、「学術コミュニケーションを専門に扱える人、求む」といったことが実際に言われだしてます。内容はいろいろで、基本的には、著作権について知っていて、そのインストラクションができる人や、DMP の教育ができる人といったものだと思うのですが、理念的には、学術コミュニケーションの専門家が研究者と一緒になって支援を拡大していったらどうだろうかということがうたわれています。

ただ、この Embedded Librarian は、欧米ですらごく一部でのサービスにとどまったので、やはり普及しな

かった要因があります。

一つは、Embedded Librarian の専門家としての位置づけが困難であることです。特に医者などに交じって、図書館員がどう専門職として位置づけられるのかということはなかなか難しいです。もう一つは、個別プロジェクトへの対応がなじまなかったということです。図書館員は無限にいるわけではないので、一人の人が幾つものプロジェクトに対応するのは難しく、かといって、あのプロジェクトには付いているけれどこちらは付けないのかという話になると、これも困るわけです。専門的なスキルや知識を持ちながら、今まで専門職として認められなかった方たちが、こういう新しい専門職としての位置づけをどうやって獲得できるのか。これはなかなか難しい問題をはらんでいます。

5-2.学術情報支援部門の構築

もう一つの方向性は、大学図書館だけではなく、さまざまな部門が連携して、広い意味での学術情報支援部門をつくっていくというものです。組織の名称上既にできている大学は幾つもあるのですが、実態として本当に協働できているかというと難しいところです。

大学全体での研究支援として、研究データ保管用プラットフォームの構築、DMP の義務化や推奨に加えて、研究成果のデータベース、オープンアクセスの状況把握、グローバルでのインパクト評価、外部資金とのリンク、今の大学図書館とビッグディール系の購読雑誌との関係をどう見るのかといったものを含めた、広い視野でさまざまな部門を連携させることが必要なのではないかと考えています。

これは図書館だけがしようとしてできることではありません。既存部門と実際に何か大きな協働プロジェクトのようなものを立ち上げてやっていくことが本当に可能なのか、これはやはりなかなか難しい話だろうと思います。

6.まとめ

このような大学の現状を考えた上で、それでもなおかつオープンサイエンス、オープンデータを中心とした学術研究支援に図書館が乗り出すのだとするならば、一体何ができるのでしょうか。

一つ、大学での個別のデータ管理・保管をどこまで進めるべきなのかということは大きな問題だと思います。現在、既に研究分野ごとのデータ公開・共有を非常に広く進めている分野はあります。ただ、それは大規模な研究プロジェクト等を中心に実現されているのであって、大学のスモールデータの集積はほとんど無視されてきています。もちろん、まずは大規模なところをきちんとやって、それをプロトタイプとして、小さいところはその後にやればいいという考えはあると思うのですが、小さいけれど、集まったときに何か意味が出る可能性があるデータを抱えているのであれば、大学は大学としてデータの管理・保管を考えてもいいのではないかということです。

ただ、どうやるかというのはなかなか難しいです。大学の個別の研究者はそれぞれの研究領域とつながって、研究領域によっては、その領域ごとにデータを管理していこうという機運があります。そこにさらに、出版社やデータ管理のさまざまなリポジトリが今できてきています。その中で、大学はどこを目指して、どのような役割を果たせるのか。大学という研究者にとって身近なところでのデータ管理・保管と、研究分野ごとのデータ公開・共有との相互補完的関係を構築し、より広い文脈のネットワークのノードのような形で機能できたら面白いのではないかと思います。

今のように、小さなデータだからというような理由で放っておいていいはずはないので、大学という場において何ができるのかという中で、図書館の役割を考えることには意味があるのではないかと考えています。

●**能勢**　大学・研究所へのアンケート結果は非常に実感しているところです。個人の感想ではなく、このように具体的な数値で示していただいたことは非常に説

得力があって、現状は私が感じているだけではなく、どこでもそうなのだなと思いながら伺いました。

途中でお話しされた、研究プロセスへの理解が必要ということは、図書館員が共通して持っている「データとは一体何か」という疑問と非常にリンクすると思うのですが、研究プロセスを理解していただくために、Embedded Librarian とまではいかなくても、もう少し簡単にできる方法のアイデアはありますか。もしくは、アンケートの中でそういったことを実行されていた方はいましたか。

これは非常に大切なところだと思っていて、研究データを実際に手にして扱うことで、研究データとは一体どのようなもので、どう管理していけばいいのかという理解に進んでいくと思うのですが、もし何かコメントがありましたらよろしくお願いします。

●**倉田**　質問調査はできるだけ答えやすいように、自由記述は一切なしで、単純で短い質問にしたつもりだったのですが、それでも回収率がこれだけ悪かったために、申し訳ないのですが、そのようなことを実行しているという回答はありませんでした。

もう 1 年前の調査なので、能勢さんの京都大学もそうだと思いますが、その後いろいろな大学において、少なくともプラットフォームの構築等に関しては幾つもプロジェクトは動いていると思います。図書館という観点からは、Embedded Librarian だけではなく、Liaison Librarian という動きもあります。要するに、研究者の方にもう少し近づこうではないかという流れは結構あると思います。

図書館員の感覚として、研究成果としての資料・論文・雑誌・図書をきちんと管理して、それを提供できればいいのだという意識が、恐らくやはり心の中にあるわけです。そこの専門家になればなるほど、それを使っている研究者をどう見るかという視点を持ちにくいですし、機会がないのです。図書館員も研究者も忙しいので、研究者が「これが欲しい」とやってきたときに、いろいろなことを聞くのは、なかなかできないと思うのです。その意味で、今、両者の関係をうまくつくることが難しくなっています。

ただ、Embedded までいかなくても、例えば今、基本的にはレファレンスにおいて、個別の研究者からの研究やデータに関する個別の要求にはほとんど応じることができないとなっているのですが、そこぐらいはもう少し緩やかに、「どんな質問でもできる限りやります」といった形を取ってもいいと思います。研究者が一体何を求めているのかは、そういうニーズをもらって、実際にそれをするのがどれだけ大変かを知る、そのやりとりの中で少しずつ分かってくる部分があるのではないかと思っています。

もう一つ、図書館員も研究すべきだと思っています。研究プロセスを知るためです。もちろん研究はそれぞれ違うわけで、私自身、全部が分かっているわけでは全くないのですが、何かの分野で何か研究をきちんとしたという経験があれば、研究プロセスを想像するのが少しはたやすくなると思うのです。また、図書館員が研究者に対応するときに、学位というのは一つの自信になるのではないかと思います。それだけでは何もできないのですが、できるところから一歩ずつという意味ではそういうこともあってもいいのではないかと考えています。

●**能勢**　ありがとうございます。Embedded Librarian や Liaison Librarian など、私の知らない言葉が出てきて、非常に勉強になりました。

●**大澤**　私もこの後話をさせていただくのですが、新しい役割の可能性として同じようなことを考えていて、Embedded Librarian については、「ああ、こういうのが既にあるんだ」と思いました。欧米では成功例がそれなりにあると伺いましたが、これはある意味で、研究者による図書館員の評価も行われるのではないかと想像しました。例えば「何々大学の何々図書館の何々さんか」と言われるような、スーパーライブラリアンというポジションができたりするなど、図書館員の個人

の技術評価の仕組みあるいは実例はあるのでしょうか。それとも、それとは違い、「あそこの大学の図書館はすごいんだな」といった形になっているのでしょうか。その辺を教えてください。

●**倉田**　私はその場で実際に見たことがあるわけではないので、あくまでも文献上の知識なのですが、医学図書館のハンドブックや教科書で、節になる程度の実績はあるということです。スーパーライブラリアンは多分いるのだと思います。図書館員の世界の中では、「あの人はすごいよね」ということはあると思いますが、どちらかというとあくまでも一つのサービスの種別として、「あそこの医学図書館にはああいうサービスがあるよね」「あの部門ではこういうことをやっているよね」と捉えられているのだと思います。

Embedded Librarian は、研究者の教育を支援することもあります。授業中にもずっといて、授業に関連する資料も全部集めてくれるなど、単なる事務サポートや TA（Teaching Assistant）ではなくて、研究者自身をサポートするということも一部の大学図書館では行われています。

●**能勢**　いかがでしょうか。皆さん、今回のセミナーのテーマは「図書館員と研究者の新たな関係」です。今、研究者 2 人が質問しました。ぜひ図書館員から、新たな関係のためにぜひよろしくお願いします。

●**服部**　SPARC Japan 事務局の服部と申します。国立情報学研究所（NII）の図書室の担当でもあるので、図書館員です。

倉田先生へというより、この会場の皆さまへの質問なのですが、倉田先生が行われた図 6 の調査の「情報管理」の記事を見ると、実際に回答された方の割合が、事務職員が 63.6%、図書館員が 12.6%でした。宛名が「研究データ管理保存関連部署」で、図書館に届いていないということが分かりました。皆さんの中で、この質問票を 1 年少し前に受け取ったことがあるという方はいらっしゃいますでしょうか。いない模様です。すみません、NII の図書室にも届きませんでした。

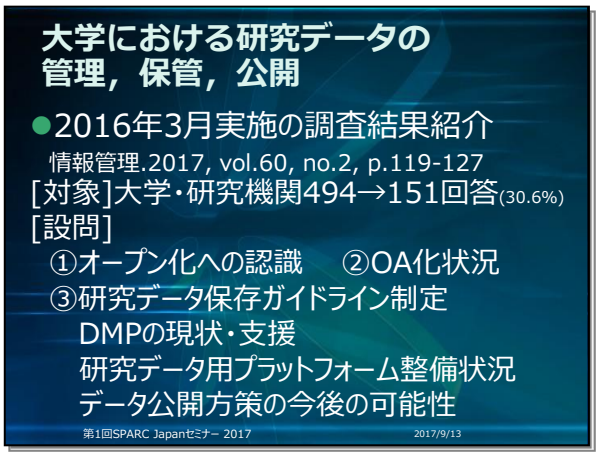
●**倉田**　申し訳ありません。どこかに行ってしまったのだと思います。

ただ、図書館員の回答も一応それなりの割合はあります。また、わざわざ、「やっと 3 カ月後に図書館に来ました。これから答えていいでしょうか」というメールを下さった方もいるので、むしろ図書館を指定して出した方がよかったのかもしれないと思っています。

●**能勢**　他、いかがでしょうか。こういうセミナーにいらっしゃった図書館員の方は、研究データを扱うことに興味を持たれている、研究データを図書館でどう扱うのかを考えていこうとされているパイオニア的な方だと思うのですが、研究プロセスを理解するため、研究データとは一体何かを知るために、研究データを扱ったことがある、または今後やってみたいという方はいらっしゃいますか。

●**林賢紀**　国際農林水産業研究センターの林賢紀と申します。図書館員をしていますが、昨年 9 月に学位を取ったので、研究者のような仕事も若干しています。SPARC Japan セミナーの企画ワーキンググループも担当しています。

今の質問の回答ですが、アウトプットの部分にだけ少し関わったことがあります。例えばアフリカで雑草



（図 6）

のデータベースをつくるというアウトプットがあるときに、データの収集に私たちは一切関わっていませんが、「これをデータベースとして公開しなければならないというミッションがあるので、ちょっと手伝ってくれ」と言われて関わりました。

その経験でお話をすると、「このデータベースを出すときに、こういうメタデータなり、メタデータの要素、語彙を使ったらもう少し検索性が上がるのではないですか」などのような話は多分できないのです。というのは、そこにもう Excel でつくられたデータと画像があって、「これを公開したい」と言われてしまっているので、与えられた材料で工夫しなければいけないからです。

「そのデータを集める前に言ってくれたら、こんな Excel っぽいのではなくて、もう少し何かできるのではないですかと言えたのに」とふつふつと湧き上がってくるところは隠して、「こういうデータをこういう感じで出すともっと検索性が上がるのではないでしょうか」というような提案はしています。研究の前のステップからご相談いただければ、もう少しいいアプローチもご提案できたかもしれないということです。

ただ、やはり倉田先生のご発表の中にあったとおり、たまたま関わったのが、最後にアウトプットをこうしたいという目的がある研究チームだったのですが、例えば論文を出すところで終わると研究チームにはなかなか関われないという問題もあります。正直、研究所にライブラリアンは自分 1 名、しかも兼務なので、約 120 人の研究者全員に関わるのは大変難しいです。これは人的リソースの問題だと思いますが。

お問い合わせいただければ応えたいのだけれども、もう少しそのプロセスの前段階から関わりたい。でも、今はアウトプットの部分だけ関わることがある。もう少し歩み寄れるところはあるかもしれないけれども、それは今日のテーマであるところの、研究者がライブラリアンをそのように使えるのだということが分からないと先に進めないのではないでしょうか。卵が先かニワトリが先かという議論になってしまいますが。こういう感じで、お答えになっていますでしょうか。

●倉田　はい、なっているのではないかと。今、私は PLOS ONE に投稿したいと思っているものがあり、原稿はできたのですがデータ公開が原則で、もちろん守っていない方もいるとは思うのですけれども、私は立場上守らないといけないだろうと思っています。データを公開できる形に整理するのは思った以上に面倒です。そのままの Excel を出したら「これは何なんだ」と言われてしまうでしょう。重要な元のデータの部分だけを残して、そこをきれいにしようとすると、ものすごく時間がかかってしまいます。

方針や政策的に、最後のアウトプットのところでデータ公開をという機運が高まりつつあるので、今はまだですが、日本の研究者も徐々にそういうことに対応していかなければいけなくなると思います。そのときに「どうしよう」で止まってしまうのではなく、図書館員にちょっとでも相談できればいいのではないでしょうか。「この分野だったらこのリポジトリがすごく楽にできますよ」「それだったら、ORCID さえ取ってあればここでできます」など、そういう情報でもありがたいわけです。その意味で何もできないということはないはずで、両者の意見が表に出てくる場があれば、もう少し進むのではないかと思います。

●能勢　まさに私も林さんと同じ意見で、林さんから頂いた意見は非常に貴重です。最後のステップでもいいですし、できれば少し前のプロセスから関わっていただいて、このようにするとデータが公開しやすくなるという情報をもらえると「そういうことだったのか」と分かりますので、研究者にとって非常にありがたい話です。素晴らしい答えをありがとうございました。

第 1 回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係:研究データの管理と流通から考える」

学術リポジトリは研究者と図書館員を繋げるのか？

大澤 剛士

(国立研究開発法人 農業・食品産業技術総合研究機構 農業環境変動研究センター)

講演要旨

言うまでもなく研究者の最も重要な職務は、研究成果を論文として公表することである。公表した論文が広く読まれ、科学の発展や社会的な課題の解決に貢献することは全ての研究者の望みである。この観点から、論文の流通を促進してくれる学術リポジトリの存在は、研究者にとって多くのメリットがある。その反面、研究以外の業務が増大している昨今、リポジトリ利用に際して新たな労力が発生することは避けたいというのも偽らざる本音である。研究者と図書館員が持続的な形で協同するためには、いずれかのみではなく、互いにとって利になる形を考えていく必要がある。本講演では、研究者の現状と研究成果の流通に求めること、さらには演者の個人的な考えに基づいた図書館員との連携の可能性について課題等の問題提起および進むべき方向性について議論を行いたい。

大澤 剛士

（研）農研機構 農業環境変動研究センター 主任研究員。博士（理学）。専門は生物多様性情報学。生物の分布情報を中心に、環境科学にかかわるさまざまな情報のデータベース化および、それらを利用した研究に取り組んでいる。世界中の生物多様性に関する情報の収集およびオープン化を進める国際的取り組みGBIF（Global Biodiversity Information Facility）日本ノードJBIF運営委員。2017年日本生態学会宮地賞受賞。

私は、今日は研究者の立場から話をします。学術リポジトリについて、そしてそれを扱っておられる図書館員の方々に私たちがどんな期待を持っているのか、あるいは私たちはどんな考えを持っているのかということをお伝えしたいと思っています。

図書館員との連携への期待と不安

私自身、図書館員の方々と話をする機会が本当になく、SPARC Japan に参加していろいろ話をする機会が欲しいなと思っていたところなので、今日この場に来られて、クラスのちょっと気になるあの子が来ている飲み会に参加できた、そんな気持ちになっています。

次はあの子からいかにして電話番号を聞き出すかがポイントになりますので、少し長めに議論の時間を取りたいと思います。皆さん、遠慮がちにではなくて、堂々と質問してきてください。質問が出なかった場合は強引に指して「どう思う？」と聞きますので、ぜひ質問等を考えておいていただければと思います。

最初に、私は SPARC Japan への参加・いろいろな方々との議論を通した図書館員との連携、そしてリポジトリを通した図書館員との連携にものすごく大きな期待を持っています。

けれども、同時にものすごく大きな不安も感じています。それは一言で言うと、図書館員への雑用・業務の押し付けになってしまうのではないかということです。

ということで本日は、図書館員と研究者のお互いにとって本当に利がある形、お互いが幸せになって楽しい形は一体何なのかについて、答えは出ないと思いますが、目指す方向、もしくは一体どこが論点になるのかを見つけられればと思っています。

もう絶対に忖度不要で、「こういうことを言ったら、ちょっと嫌な気持ちになるんじゃないかな」などは、なしにしてください。私も気にせずに思い切り妄想をぶち上げますので、ぜひ「それは全然違う」という駄目出しでも構いませんから、いろいろ意見を頂ければと思います。

自己紹介

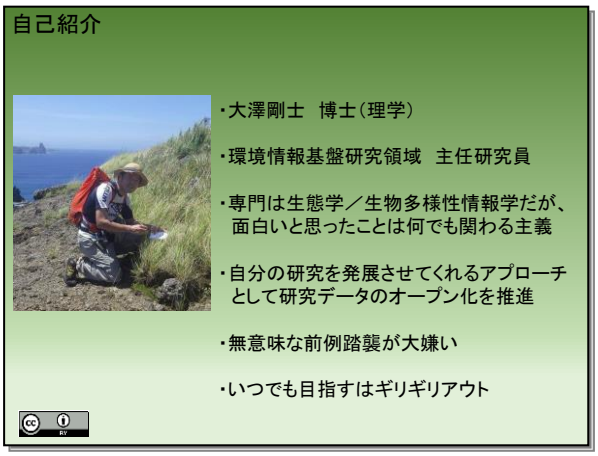
私は、農業・食品産業技術総合研究機構（NARO）農業環境変動研究センターというつくばにある研究所の、環境情報基盤研究領域で研究しています。一言で言うとデータベース部門です。けれども、私自身がデータベース、IT が専門の人間かというとそんなことはなく、もともと生態学・生物学の研究者です。図 1 はフィールド調査をしている写真です。本質としては、自然いっぱいのところで葉っぱの枚数や草の数を数えたりして、「ああ、ここは多いな」と思ったりする、少しおたくな感じの研究をしている人間です。もちろんこれは楽しいのですが、それをやっていると、やはり自分の目に見える範囲のフィールドの限られたものしか研究できないと思いました。それに対して、データベースで、他の人たちが取ったデータも全部まとめ上げることで、よりいろいろな研究ができるなということに気付き、そこに興味を持って、データベースを扱った生物学、私は「生物多様性情報学」と呼んでいる研究をしています。

このデータというものは、分野にかかわらずいろいろなものに関われる道具なので、私は最近はもはや生物学以外にも関わるようになっています。データベースは、その中に使えるデータがないと自分の研究では使えません。当たり前ですよね。ということで、自分の研究を発展させるアプローチとして、オープンデータ、研究データのオープン化というものに非常に興味を持っていて、それの推進に力を入れています。

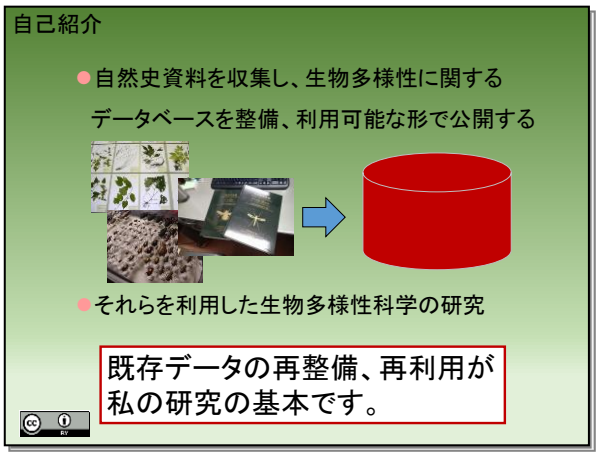
他には「無意味な前例踏襲が大嫌い」「いつでも目指すはギリギリアウト」という、うさんくさいことを書いていますが、私の話を聞いていただければ少し分かると思います。

私の研究は、生物学メインで考えた場合、基本的には標本が相手です（図 2）。ただ、標本そのものを見るという分類学者の研究ではなく、いろいろなところにあるさまざまな標本をまとめ上げてデータベースをつくり、巨大なデータセットをつくって、それを使って研究を行うというものになります。

皆さん生物学者ではないので、「えー、そんなもの



（図 1）



（図 2）

があるんだ」というぐらいの認識だと思いますが、この生物に関するデータを、データベース、巨大なデータセットをつくってみんなで共有しようというのは国際的な流れでもあります。Global Biodiversity Information Facility という一種の国際機関が、世界中の生物の情報、主に分布情報や標本等の情報データを電子化して集めて、世界中でオープンデータを誰でも自由に使えるようにしようといった取り組みを行っています（図 3)。それの日本のグループ、日本ノードと呼ばれるものがあり、私はそこの委員としてこの活動にも関わっています。

生物多様性情報学の研究

せっかくなので、少しだけ私の研究を紹介します。非常に簡単に言うと、日本全国を対象にした絶滅危惧植物の分布データをデータベース化しています。また、私は農業の研究所で働いているので、最近、農業を行わない農地が増えていますが、どんな場所が耕作放棄されているかという地図をつくっています。それを比べたら、絶滅危惧植物の分布と耕作放棄地の場所は全国的にきれいに重なっていたという結果が出ました。これは PLOS ONE に出しました。ですから、このデータは全部オープンデータになっています。

もう一つは、それと同じデータを使っているのですが、近代農業というと、機械化や大規模な農地を想像すると思います。その大規模化した農地と、絶滅危惧植物の分布は全国的にきれいに排他しているということを明らかにした研究を行いました。一人ではこんな全国レベルの調査・研究などできやしないのですが、データベースを使うとこんなことができます。

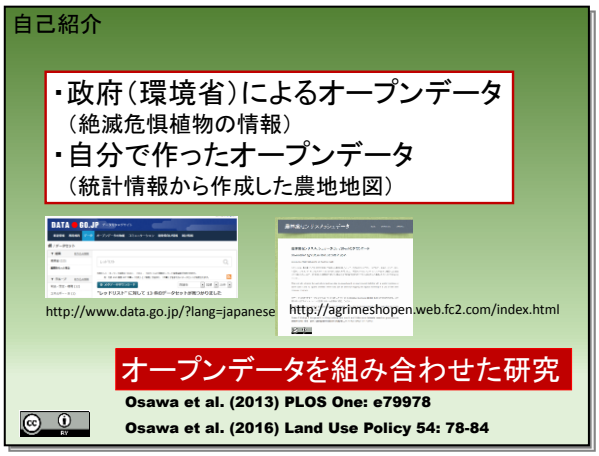
ここで今日の文脈でお話ししたいのは、これはオープンデータを使った研究だということです。絶滅危惧植物のデータを一生かけて調べてこいと言われたら私は喜んでやるところですが、現実的にはできないので、環境省が公開している、全部ではありませんが、国内で減っている植物・生物の分布データ（レッドリスト）のデータを使いました（図 4)。このデータはオープンデータとして誰でも自由に使えるようになっています。

農地、耕作放棄地のデータに関しては、農林水産省が実施している統計情報のデータを加工して、統計情報も事実上オープンデータのようなものなのですが、加工したデータをさらに改めてオープンデータとして公開しました。まさにオープンデータによった、巨大データベースを使った研究をやっているということです。

ということで、私はオープンデータというものに非常に興味を持っています（図 5)。オープンデータ憲章で取り上げられている政府情報のオープン化にも興味を持っていて、研究の分野でもこんな形でオープンデータに関わるべきではないか、こういうことができる、こういう事例があるという文章を書いたりしています。

それが現在の私の仕事にもなっていて、小さなプロ

（図 3）



（図 4）

ジェクトですが、所内でオープンデータに関する課題を立ち上げて、そこのリーダーを務めています。CKAN カタログサイトという、日本が政府情報をオープンデータとして公開しているサイトと同じ仕組みのサイトをつくって、農業データのオープン化をここで進めようとしています。

オープンデータも流通しなければ使えないので、「研究情報、データ流通は研究データのオープン化の鍵」だと言えます。私はこれが目的なのではなく、私自身の研究にとってプラスになるのでこれに取り組んでいます。

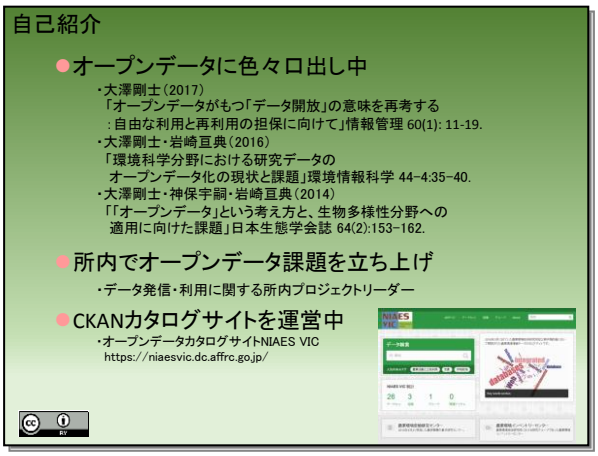
研究者の願い

少し前置きが長くなりましたが、ここからが本題です。私の研究者としての願いは、とても単純です。研究を頑張って、クールな論文を書いて、多くの人に読んでもらって、ドヤ顔しながら楽しく生きるということです。何かすごく俗な表現ですが、引用数が増えるとやはり幸せです。今、Google Scholar Citations などを使うと、自分の論文が何回どの雑誌から引用されたかをぱっと調べられるのですが、自慢できるような数ではないので少し画質は落としてあります（図 6）。当然、引用されるということは科学に貢献できたということでもあり、すごくうれしいのです。これは研究者としての素直なモチベーションです。

けれども、当然ですが、研究して論文を書いていくというのは簡単なことではありません。もちろん英語で論文を書くというのは日本人にとってすごくしんどいですし、せっかく論文を書いたから、「ぜひ引用してね」と言っても、「うちの大学からは、最近あそこの雑誌でないと取れなくなっている」などという話もよく聞きます。

それを何とかするためにオープンアクセスにしましょうという話は出ているのですが、例えば皆さんよくご存じだと思いますけれども、E が付く出版社がありますよね。そこの雑誌で普通のピアレビュー付きジャーナルをオープンアクセスにしようと思ったら、1 本、日本円で 30 万円ぐらいかかります。PLOS ONE の 2 本分、ちょっとしんどいですよね。そのお金があるのだったらともかく、研究費は激減しています。オープンアクセスにしたら研究費なんかなくなってしまいます。

そういう現状があるので、研究者用の SNS などもはやっているのですが、こういう場で話す人間として見ると、それはいろいろな問題があります。もちろんポストプリントをきちんとつくるという選択肢もあるとは思うのですが、そのポストプリントをどこに置くかを考えたときに、やはりリポジトリがあると、いろいろな人に読んでもらう機会が増えて、引用数も増えて幸せだろうと思うのです。ということで、このリポジトリを図書館で管理してくれるということは、研究者にとってメリットしかないと思っています。



（図 5）

（図 6）

研究者の悲しい現実

次に、私が研究者として避けたいことは、いわゆる雑用、謎の書類書き、はんこ押しです。こんなこと誰でもやりたくないです。そして最近話題になった「ネ申エクセルの修行」、こんなものやりたい人なんかいないだろうなと私は思っているのですが、もしかしたら好きな人もいるかもしれません。私は嫌です。

どことは言いませんが、とある私が経験した組織には、研究成果管理システムというものがあります。これを使うためには、108 ページのマニュアル、このページ数がまた嫌味が効いていていいですよね、を見て、半日かけてやり方がやっと分かったということで、ぽちぽちと頑張って登録したら、「この部分に不備があります」と事務方に怒られて、修正したら、「あまりにミスが多いのでマクロをつくりました」と登録用マクロとやらが出てきて、そのマクロを動かすために、某バージョンの Excel を入れて「ああ、やっと動いた」と思ったら終わりではなくて、それを印刷して、はんこを押して上司に提出する。そんなことを業務としてやっています。今年もやると思われます。

こんな現状がある中で、こういう言い方をするとごくネガティブに聞こえるのですけれど、さらに別立てでリポジトリというものが出てきました。それは JAIRO らしいのですが、まだ外から見ることができません。その研究成果管理システムでかなり食傷気味なところに、「これにも登録しなさい。別のものだよ。分かりやすいマニュアルもあります」と言われて、108 ページはありませんでしたが結構ごついマニュアルが出てきて、しかもそれをざっと見てみたら結構いろいろ間違いがあったのでツッコミを入れたら、その後スルーされたままで、その状態で「そろそろ登録を始めてください」と上司に言われているので、私自身も今スルーしているところです。

ここまで単なる愚痴っぽくなりましたが、何が言いたいかというと、すごく大変なのに頑張って研究をして、たくさん成果を出せば出すほど雑用が増えて、謎の作業が増えて、担当の人たちにも迷惑を掛けてしまっている。それで相手が悪くないと思っても、「また直さなきゃいけないんですか」というようなやりとりになって、だんだん担当との関係も悪化していくという現状があるということです。

ライブラリアンに聞きたいこと

以前の SPARC Japan セミナーで、研究データ、研究情報の公開などに図書館員が関われるという事例の紹介がありました。私はそれを聞いて、研究成果の流通を図書館で担ってくれるのなら、もう言うことなし、最高だと思いました。

しかし、これはいわゆる雑用なのではないかという疑問が頭をよぎりました。今日の最初の方でも、こういう仕事は、もともと伝統的な図書館員の仕事ではなかったという話がありました。自分だって仕事が増えて雑用が増えて嫌だと思っているのに、図書館員だったらオーケーなんて、そんな虫のいい話はないのではないかと、やはり少し心配するのです。

ということで、皆さんに今日聞きたいと思っているのは、ライブラリアン、司書、図書館職員にとって、研究者とより深い関係、つまり今までなかった、新たな仕事をつくるような関係を築き上げることに対して、インセンティブはあるのかということです。

もう一歩踏み込んで、私は倉田先生の講演で Embedded Librarian について聞いて、研究成果・データ流通促進を頑張ったということを、できればライブラリアン個人の業績にして、「あの図書館員さんはすごいね。ぜひうちに来てください」という話が出るような仕組みがあったらどうかという妄想をしました。これはアカデミアに身を置く研究者的な発想です。

続・研究者の悲しい現実

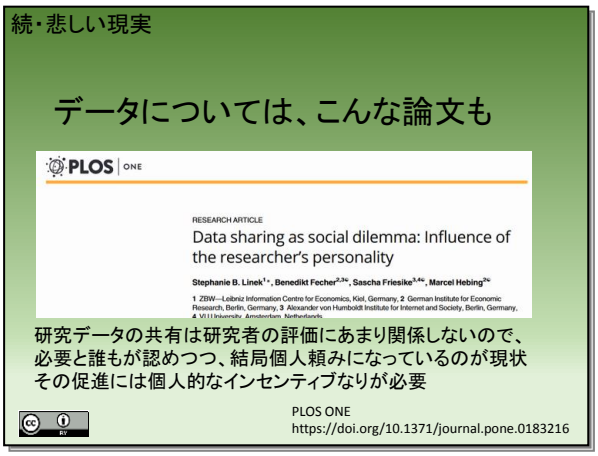
ここから研究者の悲しい現実に戻ります。きちんとリポジトリがあって、それが運用されるということは研究者にとって利点ばかりですが、実は全ての研究者にとってそうかというと、ノーと言わざるを得ません。

機関リポジトリの Wikipedia での定義は、「研究機関

がその知的生産物を電子的形態で集積し保存・公開するために設置する電子アーカイブシステム」となっています。しかし、研究所なのに知的生産を全くしない人というのも実は結構いるのです。これはまた愚痴なのですが、われわれの世代以下は、研究職に就くために苛烈な競争に勝ち抜かねばならず、何本以上のピアレビューの論文がなかったらもう面接にも呼ばれないといった状況にあるのに、その上には、管理職でもないのに、論文を書かないのか、書けないのか分からないですが、10 年間ピアレビューの論文なし、かといってデータを出している様子もなく、口は大変上手な、訳の分からない人たちがたくさんいます。

そういう人たちがいると想定した場合、考えてみてください。リポジトリがあって、その人たちに何か利点がありますか。あってもなくてもどうでもいいと思います。

そもそも、データ共有に関する図 7 のような研究を PLOS ONE で見つけました。「Data sharing as social dilemma」という恐ろしいタイトルです。「結局のところ、現状では研究データの共有など研究者の評価には全く関係しない。苛烈な競争がある中でわざわざそんなことをやる人はいない。だから必要とは誰もが認めつつも、結局個人頼みになってしまっているので、その促進には個人的なインセンティブが必要だ」という内容です。ということは、これに関わる人にもやはり個人的なインセンティブが必要なのではないかというのが、先ほどライブラリアン個人の業績にならないかということを言った背景にあります。



（図 7）

現状として、研究者が真剣に研究して、頑張って成果を出せば出すほど、雑用が増加して、その研究者をサポートしてくれる図書館員にも結局負担を掛けてしまう絶望工場になっていたら、それは論文数も減りますし、日本のランキングも下がるでしょう。

お互いの利を妄想する

でも、私が今日言いたいことはこれなのです。もう新しいアイデア、発想の転換が必要である。ここから先は私の完全なる妄想です。こういう新しいアイデアがあるということを、皆さんからいろいろ出していただけることを期待しています。

まずは、「アレな方々は置いておいて、意識高い系研究者と意識高い系図書館員だけで面白いことができないか？」です。今日こういう場に来ている時点で、ライブラリアンは意識高い系ライブラリアン、研究者も意識高い系研究者だと思うのです。ですから、まず、何ならこのコミュニティだけででも、面白いことをやって事例をつくってみないかというのが今日の私の提案です。

最後に、既存の仕組みではないけれど、こんなことができたら楽しいのではないかと意識高い系は思うだろうという妄想を幾つか挙げて、皆さんからのツッコミを待ちます。

一つ目、図書館員に研究プロジェクトのメンバーに入ってもらってしまう。Embedded Librarian の話はこれそのものだなと思いました。最近だと、科学技術振興機構（JST）などの競争的資金でも、データ・マネジメント・プラン（DMP）の作成が求められつつあります。必須ではないし、それの遂行が 100%求められるかというとそこまでいっていませんが、DMP がつくられる時代が来ています。

データ管理は、研究者にとってしんどい、面倒なところがあるのですが、基本的にやるのはメタデータづくりです。となると、私の想像ではやはり図書館員の方々が日常的にやっていることで、かなり慣れている

のではないかと思います。

内容が職務から離れないことなのでお願いしやすいということと同時に、研究プロジェクトのメンバーということは、資金が配分できるのです。そうすると、その職務に対して、一定の研究費なり管理費なりという形でお金を出すことで、リポジトリの維持やその他業務にも利用できるようになるのではないか。こういう形になったら、お互いに利があるのではないかなというのが、妄想その 1 です。

妄想その 2 は、私が昔から言っていた内容で、情報管理者（Information Manager）というポストを置くということです。これは海外では結構聞くのですが、日本には事実上存在しないと言っていいと思います。要するに、プロジェクトのデータベース、研究機関のデータベース、研究データや研究成果といったありとあらゆるデータの管理者を専門職として置くという考え方です。

私はアメリカなどでその話を聞いて絶望してきました。例えば、とある研究サイトを想像してください。そこに Associate Professor/Information Manager のようなポストがあって、その人は Ph.D.ではなくてエンジニアなのです。日本だったら何となく情報担当の准教授のような形で、何となく得意な人がくっついたり、悪い場合にはポスドクが押し付けられたりすることもあると思いますが、専門職が存在していました。「こんなのがあればいいな、こんなのをつくるべきだよ」と私は折に触れて主張していたのですが、実は図書館は、リポジトリという箱と、それを管理しているライブラリアンという管理者がそろっているわけです。それがぽんと、もしかしたら新たなロールとして収まるのではないかというそんな妄想をしました。これは業務との連携はもちろん、研究者的な業績にもなり得るので、もしかしたら先ほど述べたスーパーライブラリアンのようなステップアップにも繋がるのではないかと思います。

妄想その 3 として、もう少し踏み込むと、ライブラリアンは、もはや共同研究者、メンバーの一人だという考え方もあり得ると思います。最近はデータ論文、データを公開することに対するインセンティブとして、データそのものをピアレビューの論文として扱うという形態が出ています。そのときに、データの整理はすごく面倒です。技術も必要です。それを整理してくれる担当の方がいたら、間違いなくその研究において欠かすことができないメンバーなわけで、共著者として入ってもらうというやり方は、業績、インセンティブの積み重ねになるということもあり得るのではないかと思います。

今、私は三つ挙げましたが、単なる私の妄想です。新しい仕組みとしてこんなことができるのではないかという例であり、こういう話が皆さんから出てくればなと思っています。

いい話や、「こうするべきだ」という話はそこら中でされるのですが、もうそういうものはなしにしましょう。お互い面白いでもいいですし、お互い自分の仕事が楽になるでもいいですが、利がある形でないと、やはり持続性、継続性がありません。面白いこと、明るい未来を考えながら、残りの時間でいろいろ議論させていただければと思います。

●**林和弘**　図書館ではないのですけれども、NISTEP の林です。SPARC Japan の運営委員と、この SPARC Japan セミナー企画ワーキングの主査をしております。

図書館の方々は、大澤さんの研究者像はどれぐらい特殊なのか、普遍なのかと思っている方が多いと思うので、その辺をまずお聞かせ願いたいです。

以前 NISTEP でも異端児風にお話しいただいたのですが、こういうご自身の今の研究者像は、農研機構や生物分野の中でどのぐらい一般的なのか聞かせていただければと思うのですが。

●**大澤**　多分、マジョリティではないだろうなということは理解しています。恐らく皆さんもいろいろな研

究者と付き合っていると思うのですが、「専門はこれです」と一言で言うのが良く、あれもやる、これもやるというのは良くない、そんな空気感があるのではないかと思います。私はその逆で、何が専門と聞かれると困るぐらい、面白いと思ったことは何でもやろうと思っています。

ただし、分野も問わず何でも手を出す代わりに、中途半端にかじるだけではなく、研究者として論文にして、それを発表して、オープンにしていこうというスタンスでやっています。ですから、恐らく、一般的な研究者とは、研究スタイル、研究に対するものの考え方がだいぶ違うと思います。昔からそういう考え方を持っていたのですが、5 年、10 年とやっているうちに、少なくともこのオープンデータ、データベースを使った利用・研究に関しては、「それ、いいよね」と言ってくれる人が増えてきたという実感はあります。

それはなぜかというと、やはり利があるからなのですね。冒頭に話したように、私は昔、今でもやっていますけれど、草の数を数えたり、葉っぱの枚数を数えたりといった研究をしていたのですが、データを取るのがすごくしんどいのです。だから、いろいろなデータを使えるようになると、やはり楽な面は当然出てきます。データを管理するという別の技術は必要になってきますが、それがおいしいと皆さんが思うようになってきたという時代です。要するに、このオープンデータ、データベースをつくるということは、今、研究者全体にとって大きな興味になっているので、私の話自体は、研究者の興味から思い切り外れているというわけではないと確信しています。

●**林和弘**　今のように研究者も新しいサイエンス、データ駆動型のサイエンスに四苦八苦している中で、ライブラリアンとのいろいろな手の取り合い方があるのではないかと思うのです。

大澤先生はこういう研究プロファイルであると。では、こういう先生の研究活動で、こういうデータを扱っているときに、どういうパートナーシップがあるのかということについて、ぜひ忌憚のない意見を出していただければと思ってお聞きしました。

●**大澤**　ありがとうございます。いかがでしょう。では、申し訳ないけれど、同僚の方、よろしいでしょうか。NARO の方。

●**フロア 1**　農研機構の図書情報室の職員です。生粋の図書担当者で、職責も行政職というか事務屋ですので、研究員の方と同じ立場でお話しするという感覚はあまりなく、話も苦手な方なので何を話せばという感じなのですけれど。

農研機構の図書情報室でも、機関リポジトリを担当して構築を開始しています。こういう部門では、大学に比べて農研機構は 10 年遅れているといわれていて、研究員にこういう論文を集めるというイメージがまだあまり付いておらず、大学の皆さまの後を追い掛けて、やっと機関リポジトリができるかなというところです。ただ、機関リポジトリ用のデータを集めるときに、研究員にお手間を取らせるのもというところで、農研機構では研究成果を登録するためのデータベースというシステムが業績評価等のためにあります。その一環として、出した論文を登録していただけるだけでいいように、二つシステムはつくらずに、一つにまとめて、そこにリポジトリ用のデータもくっつけて、その結果を国立情報学研究所（NII）の JAIRO Cloud に載せて公開させていただくというような形です。昨年ちょうど農研機構も 4 法人統合して、足並みをそろえてということで少しずつ進めていたというような状況で、少しでもお手間を掛けないようにという考えの下に機関リポジトリをやっているというのが 1 点です。

また、農研機構は研究員が 1,300 人ほどいて、全国にある 21 機関の研究組織を束ねています。食品の微生物の繊毛を数える研究や、食品輸送、リモートセンシング、動物の病気の研究など、いろいろな研究をしている研究員が、北から沖縄の糸満まで全国に散っているという大きな組織です。

その中で研究員の成果をきちんと公開していくためには、どうしても単一的な一つの規律をつくって平等にやっていかなければいけません。ただ、今クリエイティブ・コモンズなども貼られていますが、実際、機関リポジトリにそういうものを書くと、その意味すらも分からないので載せてくれと言われるなど、研究員の間でも理解度の違いはあります。研究員として一生懸命頑張って、良い成果を出していただければ、事務屋としてはうれしいのですが、機構全体の研究員を見たときに、必ずしもそれが平均的ではありません。少しずつ農研機構の研究員の底上げをしていきながら、少しずつオープンデータの理解を深めていければと考えて日々仕事をしているのが私の立場です。

共著者として名前を連ねるというのは、もうそれは私ではなく研究室でどうぞお金を取って優秀な人を雇ってくださいという感じで、図書屋としては、組織として運用がきちんとうまくいくように日々努めています。

●**大澤**　農研機構は、研究成果管理システムが出てきてから、「登録ミスがあったので直してください」というやりとりを恐らく日常的にされていると思うのですが、その結果として雑用をお互い増やし合っているような感覚があるのです。そういう感覚はあまりないのですか。私見で結構ですのでお答えいただきたいです。

●**フロア 1**　それは企画の担当の事務屋が 2 名ぐらいでずっとしています。各研究所で取りまとめている方々がまた別にいらっしゃって、幾つもの手を経てやっているところもあります。始めて 1～2 年目で、システムの不備もあり、データの入れ直しはあまり自由が利かないので、できるだけ直しながらやっていくのだろうと思うのですが、何分システムを改修するとなるとお金も時間もかかります。頂いた意見は累積はしていただいていますが、反映していくのは次のステップなのではないかと思います。

●**大澤**　よろしければ、総合大学の図書館の方にお伺いしたいです。今、農研機構は微生物や畜産までいろいろな研究をしているという話がありましたが、総合大学はもっとヘテロだと思うのです。文学部から理学部、医学部、歯学部まであるわけで、文化が違う中で、リポジトリなりの情報を集めてきた中で得た経験や出てきた問題はわれわれにとってもすごく参考になると思いますし、よその方々にも参考になると思います。何かありましたら聞かせていただきたいのですが、いかがでしょうか。

●**フロア 2**　京都大学附属図書館のリポジトリ担当です。今日のお話は本当に面白かったです。

今のお話は、いろいろな分野がある中で、リポジトリにデータを集めたり、運営したりするのは大変ではないのかという話だと思います。確かに大変は大変というか、私たちもそれぞれの分野に合わせてリポジトリの論文のデータを集めるというところまではできていないといいますか、オープンにしたいという先生方の求めに応じて登録しているという状態で、なかなかそこは難しいなと思っています。

ご質問にはうまく答えられないのですが、先ほど先生から、雑用を増やしているのではないかと心配しているというお話があったかと思います。それについて、今私はリポジトリの担当なのですが、3 月までは普通の図書館員としてカウンター業務もしていました。図書館員は、先生方から何かを言われたら、それに応えてお役に立つのが仕事なので、雑用を増やされているという意識は今まで持ったことはあまりないです。他の図書館員もそういう人が多いのではないかと思っています。仕事なので、自分個人の実績になるということにあまり関心はなく、お役に立てればうれしいので、組織としてお役に立てるような協力関係ができたらいいのではないかと思いました。

また、図書館員は研究に関する専門知識が全くないので、専門知識がないのに研究データについてお役に立てるのかどうか、こういう話を聞くたびに不安に思

っています。特に、うちの大学のように総合大学になると、なおさら最先端の研究をされていて、研究内容などはっきり言って、私たちにはさっぱり分かりません。その中で、図書館員がお役に立てることはあるのか、とても心配に思っているところです。すみません、全然違う話になりましたが。

●**大澤**　いえ、すごくありがたかったです。ちなみに、今の発言に対して「私、図書館員だけど、それは違うぞ」という方は、もしかしていますか。いなかったら、私にとっては、今の話はすごくほっとするところなのですよ。情報流通の話にしても、私たちにとってはどうしても雑用的な部分がどうしてもあるのですが、それをやること自体が職務として問題ないということであれは、課題はやはり研究者側のリテラシーというか、自分の職務がどこまでかという線引きになるような気がします。私ばかりしゃべるのも何なので、研究者の方、何人かいらっしゃいましたよね。「データのリポジトリとか出すのは面倒」「ここまでだったらやれるのに」など、今の話を聞いて何か言いたいことがあるという方はいらっしゃいませんか。

PLOS ONE の話も出ましたが、最近、論文を投稿するときに、別のオープンリポジトリからのデータへのアクセスも担保する、DOI も付けるというのが当たり前になってきています。そこまでやったら、同じものを図書館員に渡すことまでを研究の一連のサイクルに入れるようにすれば、実はかなり解決できるのではないかという気がするのですが、いかがでしょうか。

●**フロア 3**　JST のバイオサイエンスデータベースセンターの職員です。私は図書館員ではなくて、先ほどのご発表ですと情報管理者というものになるのだと思います。具体的には、バイオ関係のデータベースを集めてきて、それをアーカイブして公開する仕事しています。

今のお話だと、研究者がデータをつくって、それを単に図書館の人はアップロードすればいいという流れだったと思うのですが、私はちょっと違うイメージを持っています。というのは、今までの仕事の経験上、集めてきたデータは正直言って汚いのです。私のバックグラウンドはエンジニアで、バイオの知識は研究者の方に比べればなく、データの中身は深くは分からないのですが、データの形が汚いのですよね。

例えば二つのテーブルがあって、それぞれ同じ ID が書いてあるのですが、実際にデータベースに流してみると一致しないのはなぜだろうと思い見ていくと、片方の ID は 8 桁なのに、もう一つは 7 桁になっていて、ゼロが 1 個抜けていたということがあります。それはほんの一例で、いろいろなところに矛盾がたくさんあります。

そういうものは、データの中身について深く知らなくても見れば分かるものなのですが、一つ二つではなくて結構多いのです。私としては、「論文はこんなに丁寧に書くのに、どうしてデータはこんなにいいかげんに出すのだろう」と思うぐらいで。「これはおかしくありませんか」と私が聞くと、「ああ、間違ってました。ごめんなさい」といった答えが返ってきます。むしろ、図書館員なりの視点で「これ、おかしいですね」「矛盾してますよね」というツッコミを入れるような関係になれればいいのではないかと思っています。

●**大澤**　ありがとうございます。それはすごく重要な気がしています。先ほどの京大の図書館の方もそうですが、研究者を先生のような扱いをしている空気があるのではないかという印象が個人的にはしています。でも、そうではないのですよね。研究が担当の人なのか、データ流通が担当の人なのかという、単なる職務の違いなので、上下関係はないのです。イーブンな立場だとそういう意見も言いやすくなると思うのですが、いまいち言いにくいような空気がもしかしたらあるのではないかという感覚が私にもあるのですけれど、いかがでしょう。「そんなことないよ」「実はそうなんだよ」という意見が図書館員の方から出てると、私としてはすごくうれしいのですが。

●**能勢**　私は京都大学の研究者なので、京大図書館の方のお答えに非常に勇気付けられました。私もデータの公開をしたくても、データを公開するための準備にはデータサーバーが必要であるとか、メタデータがないと公開しても意味がないとか、メタデータの作成は非常に手間が掛かるといったことがあり、そういうことをお手伝いいただけると非常にありがたいのですが、やはり雑務を押し付けているのではないかという大澤先生と共通の認識を持っていました。図書館の方はそういうものを雑務と思われないのだろうか、迷惑だと思われないのかと、やはり気になるところではあるのです。個人的に親しくしていただいている図書館の方に話をすると、「業務の一つですので喜んでやります」と同じようなお答えをされて、やはり図書館の方の意識はそうだったのだなということがよく分かりました。ぜひ京都に帰ったらご相談させていただけたらと思います。

最後に、大澤先生が考えた三つのアイデアは、やはり研究者の考え方だなと思いながら伺っていました。大澤先生は「インセンティブ」という言葉を何度か使われましたが、図書館の方にとってのインセンティブとはどういうものになるのかはやはり知りたいところです。研究プロジェクトのメンバーになってもらうという一つ目のアイデアと、共著者になってもらうという三つ目のアイデアは確かにインセンティブにはなるのですが、図書館の方にとっては、外部資金をもらっても持て余してどうしようもないよとか、共著に入っても別にうれしくないよとか、ないよりはあった方がいいのでしょうけれど、それが将来のキャリアパスに繋がるといったものではないのかなというのは気になります。そういうご意見がもしございましたら伺えるとありがたいです。

●**フロア 4**　専修大学の図書館員です。今日は貴重なお話をどうもありがとうございました。

先ほどからの議論で気になっていたことがあります。図書館で、私も機関リポジトリの PDF のデータを上げる作業や、メタデータをつくる作業などをしています。図書館にデータを渡してもらって、図書館がメタデータをつくるのが雑用とは確かに思わなくて、いろいろお仕事を頂けるのはうれしいのですが、メタデータでは、どういうメタデータをつくるのかということがあります。例えば今のリポジトリのデータでも、学術論文なのか、研究成果なのか、それとも本当にデータなのかということで、メタデータの junii2 などの規定に基づいてつくっているのです。今はデータがそういうものはあまりない状態なので、流通に適したメタデータを図書館員が最初からつくれると思われると、例えばこの話を聞いた専修大学の先生が図書館に持ってこられても、ちょっと困ってしまうのかなと思いました。

また、今のアイデアの話は、確かに研究者から見たらメリットですが、個人的に私はお金を頂いても仕方がないと思うし、研究業績の論文に共著者として挙げられても、私自身は論文発表などをすることはないので、やはりそれはメリットにならないのではないかと思うのです。具体的なお金などだと、あまりインセンティブにはならないのではないかと考えました。

逆に、先生方にリポジトリについて質問に来ていただいたり、図書館が持っているデータベースや雑誌などをこういうふうに使っているというご意見を頂いたりして、新しく始まったサービスの情報などを図書館から提供して、それを使っていただくことはメリットになるのではないかと思います。

●**能勢**　そろそろお時間となりました。大澤先生、ちょっと気になっていた相手からの答えはどうだったでしょうか。まだもう少しデートが必要かもしれません。図書館の方がどういう意識をお持ちかというご意見は私も初めて聞いたので、ぜひまた後の議論でも話をさせていただけたらと思います。

第 1 回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係:研究データの管理と流通から考える」

研究データ管理の組織的支援と図書館の役割について

西薗 由依

（鹿児島大学/JPCOAR 研究データタスクフォース）

講演要旨

オープンサイエンス推進や研究公正を主な背景として、研究プロセス全体を通した研究データの適切な管理への関心が高まっている。適切な研究データ管理を促進するには、研究者自身に加え、研究支援者が研究データ管理に関する知識やスキルを習得し、機関として研究者の支援を多方面で行っていく必要がある。JPCOAR 研究データタスクフォースでは、研究データ管理に関する学習教材の開発に取り組んでいる。本発表では、この教材とその活用方法について紹介するとともに、これからの図書館員が担う役割の可能性について述べる。

西薗 由依

鹿児島大学附属図書館に勤務。2017年度オープンアクセスリポジトリ推進協会（JPCOAR）作業部会員（研究データタスクフォース）。

私は鹿児島大学附属図書館で、主に図書館間の相互利用業務、レファレンス業務、情報リテラシー教育業務に当たっています。学内のリポジトリ業務は以前担当しており、離れて 5 年ほどたっていますが、学外のリポジトリ関連の活動に参加しています。本日はそのオープンアクセスリポジトリ推進協会（JPCOAR）の研究データタスクフォースの活動について、主にお話しします。

オープンアクセスリポジトリ推進協会（JPCOAR）

まず、JPCOAR について簡単にご紹介します。正式名称は「オープンアクセスリポジトリ推進協会」といい、日本における機関リポジトリを振興・相互支援することを目的とし、国公私立大学図書館協力委員会と国立情報学研究所の間の連携・協力協定に基づいて、平成 28 年 7 月に設立された新しいコミュニティです。

現在のところ、516 機関が参加しています。世界規模で研究論文、研究データのオープン化が推進されており、機関リポジトリの果たす役割にも大きな期待が寄せられる中、この JPCOAR は国内の大学・研究機関が力を結集し、各機関リポジトリの発展に向けた課題の解決に共同で取り組む拠点であると位置付けられています。

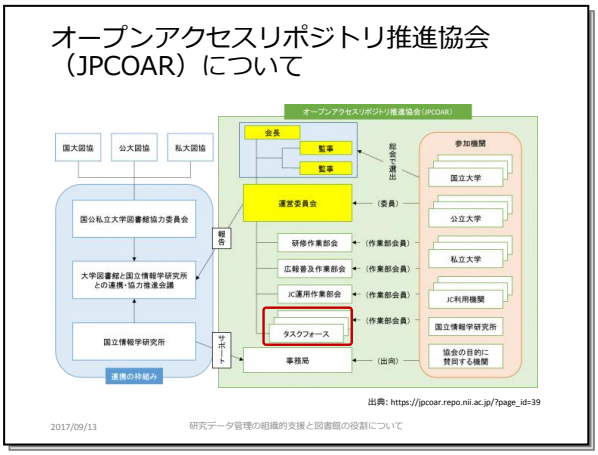
図 1 は、その課題に取り組むための JPCOAR の組織体制図です。運営委員会の下に、オープンアクセス基盤の安定的な運用を担う作業部会と、先端的な機能

の開発を担うタスクフォースが設置されています。作業部会は現在のところ三つ、タスクフォースは現在のところ四つあり、その一つが研究データタスクフォースです（図 2)。

この研究データタスクフォースは、JPCOAR の前身の一つである「機関リポジトリ推進委員会」のときから枠組みとして存在しており、活動内容も継続されています。

国内外で研究成果のオープン化が推進され、論文だけではなく、エビデンスとなる研究データの保存や共有化が求められつつある中、研究データの公開を進めるための前提となるのが、その研究データを研究プロセスの最初から最後まで適切に保管、管理することです。そのことについて学べる教材が、研究者にとっても、研究を支援する側にとっても必要ではないかということで、まず、2015 年度に海外における既存の教材の比較・調査を行い、2016 年度から実際に、研究データ管理についての基礎を学べる日本向けの教材開発に取り組んでまいりました。

教材開発に当たっての基本的なスタンスとして、オープンサイエンスを推進していくには、研究者任せにするのではなく、大学等の機関が研究者に対して組織的に支援を行っていくことが必要であると考えており、組織的な支援のために具体的にどのようなアクションを取っていけばよいのか、その手掛かりとしてもらえることを意図した構成となっています。



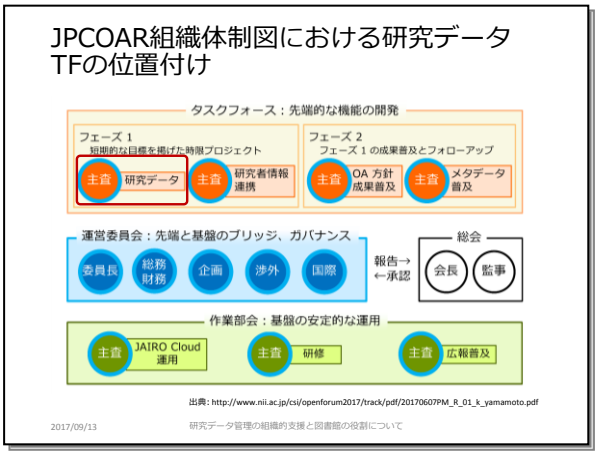
（図 1）

研究データ管理に関する組織的な取り組みの現状

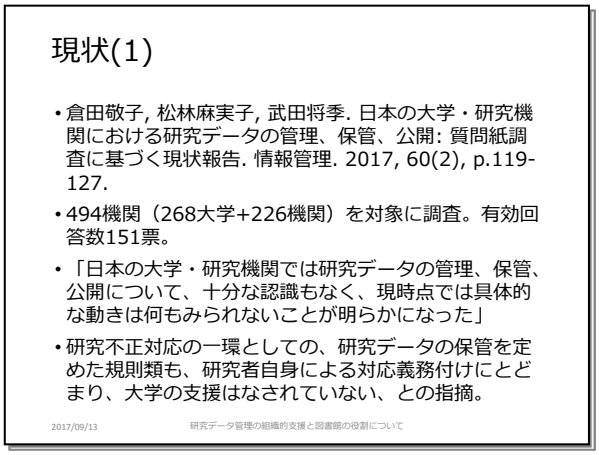
国内機関では、今のところ組織的な支援はなかなか進んでいないというのが大まかな共通認識かと思います。倉田先生の講演にもあったように、2016 年 3 月に行われた調査の報告では、日本の大学・研究機関では、研究データの管理・保管・公開について十分な認識もなく、現時点では具体的な動きは何も見られないとの指摘がなされています（図 3)。

また、研究不正対策の一環として、研究データの保管を定めた規則類の整備は、比較的多くの機関でなされてはいるものの、保管のための環境整備等、組織的な支援がなされているかというと、ほとんどなされていないという現状が報告されています。

同じような報告があります。2016 年 12 月から 2017 年 1 月にかけて、国立大学図書館協会オープンアクセス委員会が、国立大学図書館協会の加盟館を対象とし



（図 2）



（図 3）

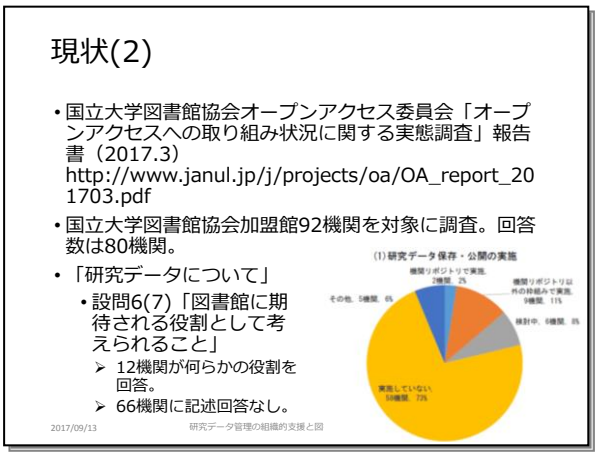
て、オープンアクセスへの取り組み状況に関する実態調査を実施しています（図 4）。その報告書でも、研究データの保存や公開に取り組んでいる機関は非常に少数であるということが示されています。

さらに、設問の中で、図書館に期待される役割について記述回答を求めるものがあったのですが、それに対しても無回答であった機関が非常に多く、報告書ではこの結果について、図書館の役割を想定できる状況にもなっていないとも見て取れる、と評しています。

国の政策としては、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」報告書をはじめ、この数年、大学や図書館に対してオープンサイエンス推進における一定の役割への期待が示されており、組織的な取り組みが求められてはいるものの、現状はまだまだこれからというところかと思います。

文部科学省の「学術情報のオープン化の推進について（審議まとめ）」では、大学等に期待される取り組みとして幾つか具体的な言及がなされています。その中の一つに、「技術職員、URA 及び大学図書館職員等を中心としたデータ管理体制を構築し、研究者への支援に資する」とあります。しかし、では、具体的に何をしていけばいいのかということについて、研究支援をする側もまだまだ認識が追い付いていないというのが現状であるように思います。

従来は、研究プロセスの最終段階への関与、つまり、出来上がった研究の最終成果物、主に論文を受け入れて、メタデータを整備して、管理・保存して流通させることが主でしたが、研究プロセスの初期段階から関与して、日々生み出される中間成果物をどう取り扱っていくのかということは、まだ経験もないですし、なかなか想像しようにも分野ごとに違うということもあり、図書館にとってはハードルが高いというのが現実のところかと思います。



（図 4）

さまざまな研究データ管理サービス

研究データ管理の組織的な支援を検討する前提として、研究者の支援に関わる職員が、研究データがどのように生成されて、研究プロセスにおいてどのように用いられるか理解を深めて、その管理の在り方を考えていく必要があります。難しいという話もありますが、そこは何か一つ進んでいかなければいけないところだと思います。

オープンサイエンス政策や研究不正対策、研究助成機関のポリシーを受けて、海外では既に大学等の研究機関が、研究データ管理サービスを自機関の研究者に提供していることが少なくありません。機関による組織的な支援として、図書館に限らず、研究を支援するスタッフが協働して、研究者によって生み出される研究データを適切に管理できるように、それぞれの段階に応じて必要なサービスを提供している事例が多く見られます。日本においても、データ管理体制を構築するために、研究データ管理に関するサービスを整えていくことがやはり必要だと思います。

研究データ管理サービスには、さまざまな種類があります。よく思い浮かぶものとしては、研究データを公開するデータリポジトリ、リポジトリ登録に際してのメタデータ付与が挙げられます。それだけではなく、アクティブな研究データを保存するストレージなどの管理基盤の提供、研究助成機関が提出を義務付けつつあるデータ管理計画（DMP）の作成支援なども含んでいます。

図 5 はタスクフォースで作成した教材で用いた図です。研究者がプロジェクトの研究助成金を得るまでのプレアワード、研究プロジェクトの実施、研究によっ

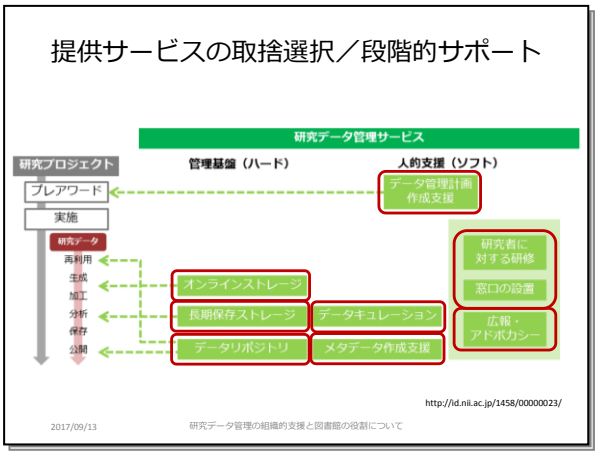
て生み出された研究データの公開に至る流れに沿って、各種の研究データ管理サービスをマッピングしたものです。

研究データライフサイクルの特定の段階で必要とされるサービスもあれば、ガイダンスや相談窓口の設置のように、ライフサイクルを通してずっと必要となるサービスもあります。このように多くの要素があるのですが、必ずしも全ての機関がこれら全てのサービスを提供しなければならないというわけではないと考えています。自分の機関の状況、あるいは外的環境の状況に合わせて、必要に応じてこれらのサービスを取捨選択する、あるいは、スモールスタートの後少しずつ段階的に提供していくということも考えられるかと思います。

例えば、まずはオープンサイエンス推進における研究データ管理の重要性について、特に若手の研究者に向けて意識喚起することから始めて、徐々に機関リポジトリでスモールデータから登録を始めてみる、ガイダンスを実施する、DMP の作成支援を行うといった進め方も一つの進め方だと思います。

国内で研究データ管理サービスを全面的に提供しているところは現時点では見られないと思いますが、研究データ保存環境の整備や、データ管理の手引きの作成・提供、研究データ管理の情報をまとめたポータルページの作成・提供など、部分的にサポートに乗り出すところが出てきている状況ではないかとみています（図 6）。

現状、日本では組織的な支援はまだまだですが、4 月に科学技術振興機関（JST）がオープンサイエンス方針を発表しましたし、出版社のデータ共有ポリシーも増えてきているので、研究データ管理の組織的な支援という課題は、日本でも重要性を増しつつあると考えていいと思います。より多くの機関において、研究支援者側が、研究データ管理に関する基礎知識やスキルを習得して、サービス構築に向けて第一歩を踏み出すこと、また、その際に図書館員だけではなく、研究データ管理に関わる可能性のあるさまざまな部署・職種の人々が共通認識を持って協働していくこと、これらが必要だと考えられます。

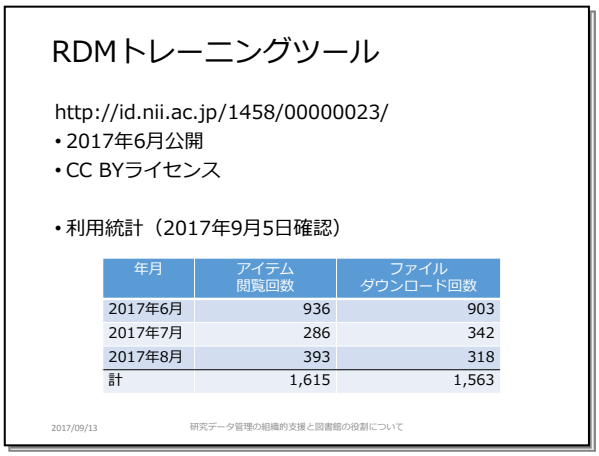


（図 5）

（図 6）

RDM トレーニングツールの紹介

このようなことのために、私たちが開発している教材を少しでもご活用いただければと考えています。私



（図 7）

たちは教材を RDM トレーニングツールと称して、6 月に JPCOAR のウェブサイトで公開しています（図 7)。利用統計を見る限りでは、まずまずのご関心を頂いていると考えています。こちらは CC BY ライセンスで公開しているので、教材として利用していただくのはもちろん、各利用者はこれを改訂したり、拡充したりといった形で自由に活用できるので、そういう使い方も期待しています。

教材の概要を紹介します（図 8)。受講者は研究支援に関わっている方で、図書館だけではなく、さまざまな部署の方をメインターゲットにしています。機関内の研究データ管理サービス構築に携わる可能性のある方々に、一緒に共通認識を持って協働していただきたいと考えています。

目的は、学習者が研究データ管理に関する基礎的な知識を習得すること、各機関における研究データ管理サービス構築の足掛かりを得ることという 2 点を掲げています。

研究支援職にとって教材というものは、自分自身の知識やスキルを習得するためのものでもあるのですが、それと同時に、それを基に、自機関の研究者への教育活動を研究支援職自身が担っていかなければならないと考えています。

教材の体裁は、講義スライドにスクリプトを付与した非常にシンプルなものです。

「第 1 章　導入」では、なぜ研究データ管理に対する関心が高まっているのか、なぜ機関として対応していく必要があるのかという背景の説明や、研究データや研究データ管理とはどのようなものかという基本的な概念整理から始めます。

「第 2 章　データ管理計画（DMP)」では、DMP の構成要素や作成に活用できるツールを紹介します。

「第 3 章　保存と共有」では、保存や共有において考えるべき留意点について学びます。

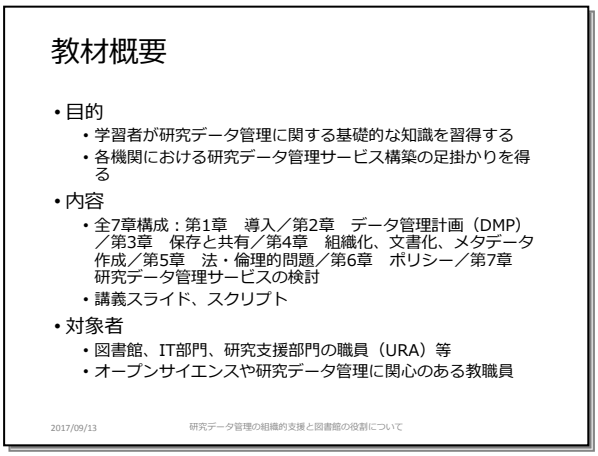
「第 4 章　組織化、文書化、メタデータ作成」では、一定のルールにのっとったデータの組織化や、データについて説明する文章やメタデータの作成について学びます。

「第 5 章　法・倫理的問題」では、研究データを取り扱う上で必要となる知的財産権、ライセンス付与について学びます。

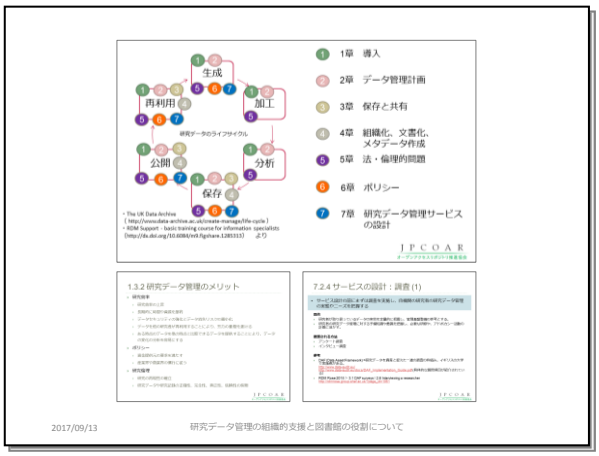
「第 6 章　ポリシー」では、さまざまなポリシー策定の動向や、各機関においても検討すべきポリシー関連のことについて学びます。

「第 7 章　研究データ管理サービスの検討」では、最終的に自分の機関の状況に合わせて、研究データ管理サービスをデザインして提供していくための具体的なステップを学びます。

図 9 はスライドの例です。教材で研究データを適切に管理して活用していくための基礎知識をまずひととおり付けていただいた上で、学んで終わりではなく、実際に自分の機関でサービスをどう設計していくか、自分の頭で考えることが重要だと考えています。一つの解があるわけではないので、実情に即して自分の機



（図 8）



（図 9）

関に合わせた独自のサービス設計を行っていく必要があると思います。

そのためにも、7 章で紹介している方法を活用して、自分の機関における研究者の研究データ管理が実際どうなっているのか、どういうニーズがあるのか、顕在的なものも、潜在的なものも含めて、その把握をする、まず実情を知るというところから始めていただきたいと思っています。

無料オンライン講座の開講

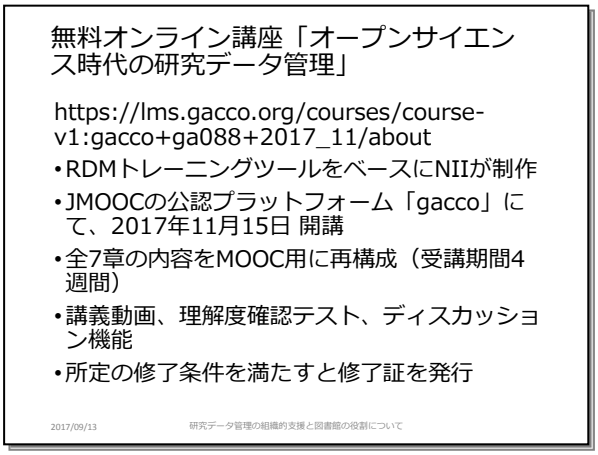
このトレーニングツールをベースにした内容の無料オンライン講座「オープンサイエンス時代の研究データ管理」も 2017 年 11 月に開講される予定です（図 10）。MOOC 用に再構成して、こちらは動画で提供しています。併せて、理解度を確認するテストやディスカッション機能も付いているので、自学自習するだけではなく、実施する上での疑問点や不安点を同じ受講者と議論する場としても、ぜひ活用していただきたいと思っています。所定の修了条件を満たすと、gacco から修了証も発行されることになっています。

図 11 は、教材開発に当たったメンバーです。

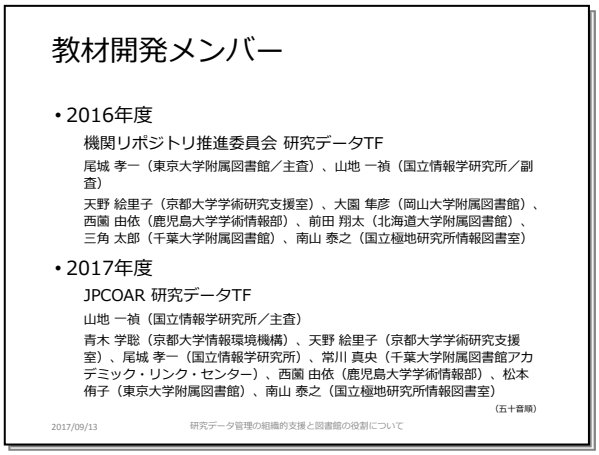
ご紹介してきた教材の活用方法としては、まずは、図書館員をはじめとする研究支援に当たる方々に、自分の知識習得のために使っていただく、機関内の研究者の方々への教育活動のために使っていただくというのが主だと考えています（図 12）。

研究データの利活用促進のために図書館が実際にどんなことができるのかという議論においては慎重な意見もよく聞かれるのですが、将来的には、この教材での学習を終えていることが、研究データ取り扱いの基礎スキルを習得していることを証明するものになるといいのではないかとも考えています。そのために、もちろん教材の改訂や拡充も必要だと考えています。

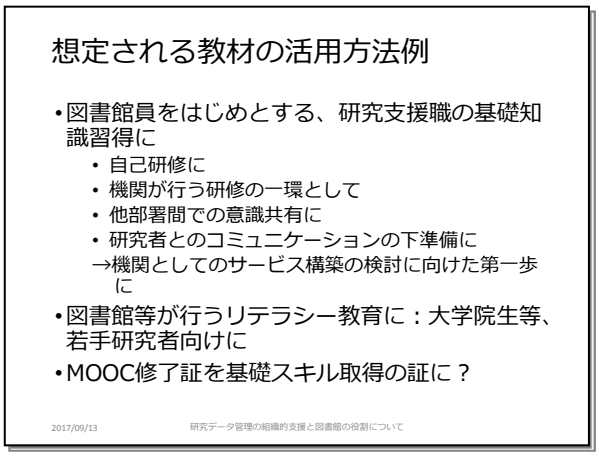
この教材は現在のところ研究支援職に向けた構成となっているので、研究者から見てどういう使い道があるか考えてみました。例えば、研究室に配属されたばかりの学生や、研究活動をこれから始めようとしている段階の方に、一般的な基礎知識を得る補助教材のような形で使っていただく、研究支援職と研究者が共通認識を持つための参考資料として使っていただく、研究データ管理に関して組織内の環境を整備してほしいという要望を出す上での参考資料として使っていただくということもできるかもしれません（図 13）。



（図 10）



（図 11）



（図 12）

いずれにしても、今回作成した教材は決して完成形ではないので、活用・参照いただいた方からのフィードバックや、各所での教材の改訂やその共有によって、より良いものに発展させることで、人材育成の土壌形成に少しでもつながっていくのではないかと期待しています。

また、教材を一つのきっかけにして、研究室と研究支援職の相互理解が進むことで、実際に各機関での効果的な研究データ管理サービスが展開されるようになり、多様なモデルの比較検討までできるようになれば、一つの成功なのではないかと思っています。

研究データ管理サービスと図書館

この教材は、図書館だけではなく、機関内の関係部署が一緒に協働していくことを想定しています。その中で、今日いらっしゃっている方々は、特に図書館が何をできるのだろうかということを気にされているのだと思います。

各機関の実情や考え方によって異なってくるのはもちろんなのですが、あくまで個人的な考えとしては、データリポジトリの提供、データへのメタデータ付与、機関内で提供できるインフラがない場合の機関外のインフラの紹介、若手へのデータ共有のメリットやデータを取り扱う上での注意点などのガイダンス、DMP作成時に各助成団体の要件と自機関で提供できる環境を比較した上でのアドバイスなどといったことは、現在行っている業務との親和性があるのではないかと思います。楽観的と言われるかもしれませんが、一定程度の貢献はできるのではないかと私は考えています。

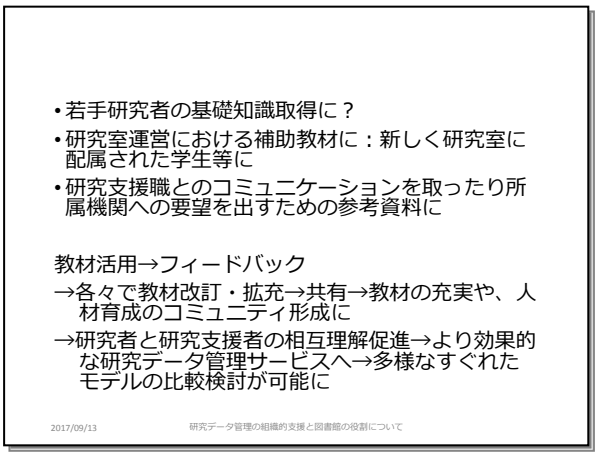
研究者の方々には、気兼ねすることなく図書館にご相談いただきたいと思っています。それが「雑用」なのかどうかは内容や捉え方次第かもしれませんが、図書館員にとって、利用者の役に立つことは無上の喜びです。基本的に何かの形でお役に立ちたいと思っているので、まずは「ちょっと聞いてみようかな」と、気軽に聞いていただければと思います。

図 14 は、昨年発表された「国立大学図書館協会ビジョン 2020」で示された大学図書館の基本理念です。各図書館は、このような基本理念に基づきながら、具体的に何を図書館の役割として果たすべきかを設定し、何をどう実現していくかを自らきちんと考えて、状況の変化に応じながら再定義していくのです。

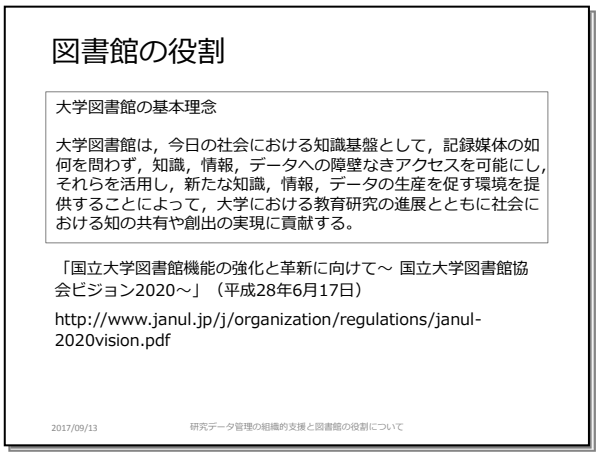
昨今、大学における教育の在り方の変化に伴い、図書館が行う学習支援も変化を求められ、各図書館は学習支援機能を強化されていると思うのですが、それと同じような文脈で、研究の在り方も非常に変化している中で、研究支援機能についてもそれがどうあるべきかを考えるべきときなのだろうと考えています。

正解は一つだけあるわけではないので、研究者と協働しながら、より良い形を考えていきたいです。まずはディスカッションすることから始めていければと思っています。

最後に、先ほどご紹介した無料オンライン講座は、11 月 15 日開講に向けて、現在受講者を募集していま



（図 13）



（図 14）

す。興味のある方はご登録いただいたり、あるいは周りの方にお声掛けいただければと思います。

●**能勢**　西薗さん、どうもありがとうございました。

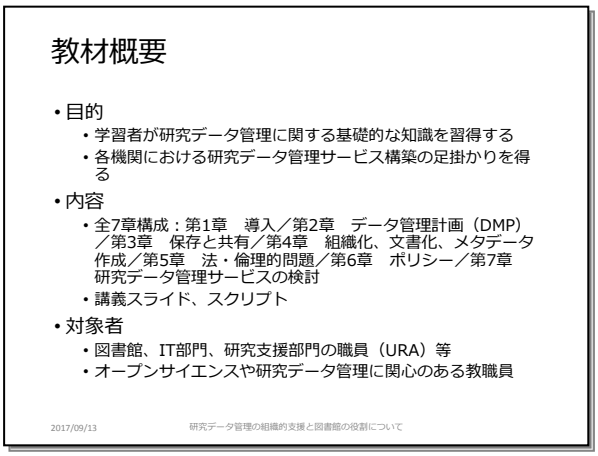
それでは、ご質問やコメントがありましたらよろしくお願いします。では、どうぞ。

●**服部**　SPARC Japan 事務局の服部です。無料オンライン講座の申し込みは結構な数来ていると聞いていますが、申し込んだ方の属性は図書館員ばかりなのでしょうか。図 8 の「対象者」に、「図書館、IT 部門、URA 等」「オープンサイエンスや研究データ管理に関心のある教職員」とありますが、そのような方から申し込みが来ているのでしょうか。

●**西薗**　私は属性までは今のところ把握しておりませんが、どなたか分かる方、会場にいらっしゃいますか。尾城さん、お願いします。

●**尾城**　国立情報学研究所のオープンサイエンス基盤研究センターの尾城と申します。この教材の開発と、オンライン講座開講の準備で西薗さんと一緒に仕事をしてきました。

まず、今のところ約 1,000 名から申し込みいただいています。私たちは 2,000 名を目標にしていて、このところ少し登録のペースが落ちてきてしまっているので、ぜひ今日お集まりの皆さん、宣伝にご協力いただければと思います。



（図 8）

今のところ属性情報は取れていません。登録してくれた方のプライバシーとセンシティブなデータに関わってくるので、どこまで公開できるかは分からないのですが、何らかの形で集計して最終的にはまとめたいとは思っています。どこまでのレベルで公開できるかは今は明言できないのですが、少し整理して今後の教材開発に活用できればと思っています。

●**能勢**　他にございますか。ここは研究者の方に発言していただいた方がいいかもしれません。研究者の方、手を挙げていただいてよろしいですか。その中で、JPCOAR が発足したのをご存じの方はいらっしゃいますか。図書館の方はもちろん JPCOAR をご存じだと思うのですが。

●**逸村**　筑波大学の逸村裕です。先ほど倉田先生、大澤先生の講演での議論もありましたが、データキュレーションとメタデータ作成支援で、取りあえずどの程度のことを今お考えになっていらっしゃいますか。

●**西薗**　本当に個人の考えになりますが、メタデータ作成支援に関しては、各機関の機関リポジトリにデータを登録してみたいという要望があった場合に、図書館員がゼロからつくるケースや、あるいは、そのデータについての最低限の情報を研究者の方に頂いた上で、図書館員が流通に適したデータを付与しメタデータを充実させるというケースを想定しています。

データキュレーションに関しては、この言葉の定義自体が人によって違うかもしれないのですが、全てのデータを永久的に保存することは現実的ではないと思うので、どれを本当に長期に保存していくべきなのかを選定する上でどういう判断基準があるのかを一緒に考えるといったことも一つの方法ではないかと思っています。

●**逸村**　いろいろなことを考えなくてはいけないことだと思うのです。例えば、海外との国際共同研究においては、日本語と英語が対応するのも結構難しいです。例えば、Digital Scholarship という言葉をどう訳すかもあちこちで話題になっています。メタデータをつくるには定義が明確になっていないと厳しいので、その辺もぜひ、いろいろなケースがあるかと思うので、そういうものを蓄積されて良いものにしていただければと思います。最後はコメントです。どうもありがとうございます。

●**能勢**　今おっしゃったメタデータの定義というのはスキーマのことになるのでしょうか。次の話で恐らくそういう話が期待できると思います。

先ほど、JPCOAR をご存じの研究者は手を挙げてと言いましたが、私は恥ずかしながら、JPCOAR のことはこの企画をするまで知りませんでした。申し訳ありません。

例えば図 12 に「研究者とのコミュニケーションの下準備に」とあるように、図書館側からは秋波を送っておられたのではないかと思うのですが、研究者で、教材開発がされている、教材が公開されているといったことを、私の身の回りだけかもしれませんが、知っている人があまりいません。それはなぜかなと聞きながら思いました。また、タスクフォースの中にどれぐらい研究者として意見が言えるような人が入っているのか、この教材はかなり図書館の方で使われることを想定したものなのかを伺いたいです。

●**西薗**　教材が全ての各分野において存在するかというと、データリポジトリの設置状況にも表れているように、必ずしも一律ではないと思います。

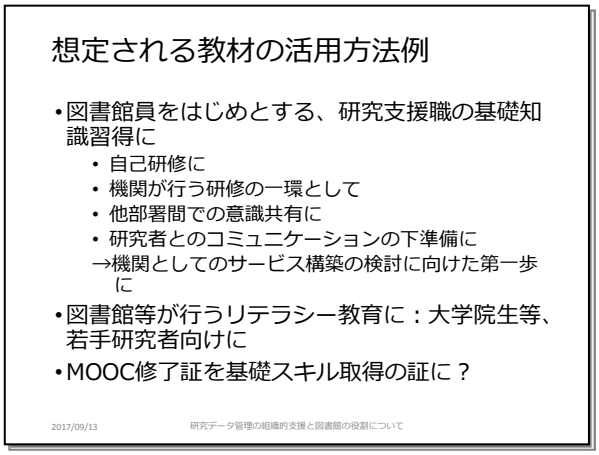
これから本当に図書館が研究データ管理に乗り出していくのであれば、特に総合大学ではいろいろな分野に目を配る必要があると思いますし、今回は研究支援職に向けた教材ですが、これをもっと研究者にダイレクトに使ってもらえるような教材開発を考えなければならないと考えています。

ただ、その上で今回のこの開発メンバーは、ほとんどが研究支援職の者なので、そこに専門知識を持った方々に参画していただく必要がどうしてもあると思います。こういうものが自分の分野にあるといいなと思っていた方、自分が使うだけではなく、学生を指導する上であったらいいなと思う方がもしいらっしゃいましたら、ご一緒につくっていくこともできるかと思いますので、ぜひお声掛けください。それは一つの協働ではないかと思います。

●**能勢**　研究者の方、いかがでしょうか。

●**深貝**　横浜国立大学の深貝です。SPARC Japan の運営委員をしていて、一応研究者ではあるのですが、このごろ、行き掛かりがあって図書館関係に親しい方々ができて、勉強させていただいています。

教材プログラムを研究者が使う場合、フィールドによってもちろんいろいろ違うのですが、それを全部入れようとすると訳が分からないことになります。結局、機関リポジトリでデータを登録したものを使う場合でも、何かの目的がある以上は、さまざまなフィールドで目的に応じたリポジトリの中からデータを拾い出すプラットフォームを工夫していくと。研究者の側、例えば学会、新しいトピックスのリサーチグループなどが、横断的にデータを集めるような仕組みをつくることを促していくのがいいのではないかと思うのです。



（図 12）

それは一挙にやると言っても無理で、誰でも使えるような横断型の擬似的リポジトリのつくり方の指南方法を図書館スタッフでつくれと言っても無理で、熱心な人をそそのかして、応援してつくってもらうというサンプルをつくるしかないだろうというのが一つです。

それから、今、西薗さんが、研究者に使ってほしい、教育にも使ってほしいとおっしゃったので、そうだとすると、これもまたサンプルになるプログラムをどこかで見つけ出す方法がいいと思うのです。私は旧式人間ですから、基本的には印刷物で、しかも古い本を使うのですが、最近の学生はそうではなく、ゼミのテキストを決めるにしても端末で調べるのです。まず図書館に行って現物を見ようとはしないのです。でも、ネットで調べても、これが良い、これが悪いというのが結構当たっているのです。そうすると、今までの学術でどうなっていたかの延長線で考えるだけではなく、新しいデジタル型人間、ネットワーク型人間たちが今後どう振る舞うのかを見ながら、観察的にスキームを組んでいくのが有益だと思います。

例えば私は、図書館の本とネットワークのデータを組み合わせて、共同作業でレポートを書きなさいという課題を学生に課しています。それを通して学生は、研究者の共同研究や、会社のチームでの仕事で責任を取り合って協力するということを学んでいます。ネットワークを使って、図書館のガイダンスを受けながらレポートを書いていく、そういうことに教材を使えるようにすることを念頭に置くといいのではないかと思います。

尾城さんの先ほどのお話では、登録メンバーが1,000 人を超えていて、属性が分からないというのは、もちろん非常にパーソナルなことでセンシティブですから、難しい面はあるのですが、それでもなお、どういう人がどういう反応をしたかをある程度匿名化した上で分かるようにしていった方がいいと思います。その理解の仕方、活用の仕方が人によってどう違うかを観察することで、今後の教材開発のストラテジーを組むことができるのではないかというのが感想です。

●**西薗**　貴重なご意見ありがとうございます。大変参考になります。得られたデータはぜひ今後の教材開発等に生かしていきたいと思っています。ありがとうございます。

●**林和弘**　林です。深貝先生とほぼ同じことを加えるのですが、今日大澤さんの講演後の質疑応答で言い切れなかったことは、大澤さんはこれまでにないデータを組み合わせたサイエンスを生み出そうとしている、そういう人がデータサポーターを求めているということです。この教材の発展の方向としては、新しいサイエンスの研究スタイルの在り方を模索し、新しい価値を生み出すという高い目標も持つことだと思います。既存の研究のある部分は大変だろうから先生のお手伝いをするというのは、それはそれで一つのミッションだと思うのですが、DMP が加わるということは、いつ、どこにどんなデータができて、それがどういう価値を持つのか、どういう共有を含む配布をすればいいのかという議論が内包されているので、新しいサイエンスのスタイルまで能動的に考えていただければと思い、余計なお世話ながらエールを追加させていただきました。

●**西薗**　ありがとうございます。

●**フロア 1**　神奈川県立生命の星・地球博物館の職員です。11 月に開講する無料オンライン講座「オープンサイエンス時代の研究データ管理」というタイトルには、「オープンサイエンス」という言葉が入っていました。この講座は、図書館員以外、例えば、私が所属しているのは博物館ですが、博物館の人に紹介することは問題ないのでしょうか。また、タイトルにある「オープンサイエンス」とはどういうことを想定されているのですか。

●**西薗**　教材の利用自体は本当に興味のある方なら、どなたでも歓迎しています。博物館も少し重なる部分

もあると思いますので、活用できる部分がありましたらぜひご活用いただきたいと思います。

オープンサイエンスという言葉の定義自体もいろいろあるのは皆さんご存じかと思います。この教材が開発された背景には、公的資金を用いた研究成果のオープン化の流れに伴い、研究データ管理が求められているということがあります。

ですので、論文やデータだけではなく、それを生み出すためのソフトウエアやコードも含めてみんなで共有していくのが目指す方向だという意味のオープンサイエンスではないかと個人的には思っています。

●**能勢**　7 章構成ですので、第 8 章として、実際にどういう協働ができたかという実例を、例 1、例 2 という形で組み入れると、他の方も、こういう協力の仕方があるのかということがクリアに分かると思います。そういうことを目標にしていただけると、非常にありがたいと思いました。

●**フロア 1**　先ほど質問させていただいた者です。先ほどお答えいただいたオープンサイエンスについてのコメントです。

皆さんご存じのように、学習指導要領の改訂で、小中学校などの調べ学習でもインターネットを使うところが増えてきています。体系的な知のコンテンツが大学や研究機関以外の場でも求められている中で、実際にインターネット上にそのようなものが日本語で存在するかというと、なかなかない状態です。

日本での知の流通の大本は研究機関のリポジトリにあると思うので、オープンサイエンスの推進といった場合に、各機関の事情や予算体系に合わせた、各機関の所属員に対するサービスとしてのリポジトリが第一にあって然るべきだと思うのです。オープンサイエンスは、ぜひ一般の方々に対する学術データの環流も想定した形で発展していただければと思いました。

●**西薗**　貴重なご意見、ありがとうございます。

第 1 回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係:研究データの管理と流通から考える」

新たな学術情報流通において JPCOAR スキーマが果たす役割

片岡 朋子

（お茶の水女子大学/JPCOAR メタデータ普及タスクフォース）

講演要旨

近年、公的研究助成を受けた学術成果および関連する研究データの公開が助成団体や政府の方針として推進されている。公開された学術成果の発見や利活用を促進するためには、研究データ等の新たな学術資源や国際的な流通に対応した標準的なメタデータ交換フォーマットが必要となる。JPCOAR ではこうした社会的要請に対応するため、機関リポジトリのメタデータ交換フォーマットである junii2 を改訂し、JPCOAR スキーマを策定した。本発表では、新たな学術情報流通において JPCOAR スキーマが果たす役割について紹介する。

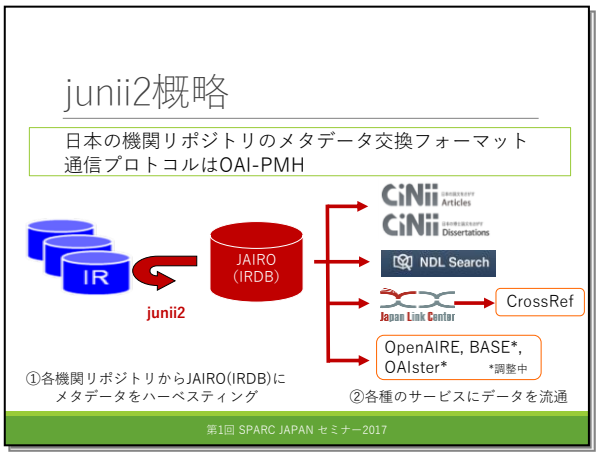
片岡 朋子

2015年よりお茶の水女子大学図書・情報課に勤務。情報基盤担当として大学全体の事務システムの運用支援を担当。2016年よりJPCOARスキーマの策定に携わっている。

本日は、私が所属している JPCOAR メタデータ普及タスクフォースが策定を進めている、新しいメタデータフォーマットである JPCOAR スキーマと、これからの学術情報流通において、この新しいメタデータフォーマットがどのように活用できるのかという事例についてご紹介します。

1.JPCOAR スキーマの概要

JPCOAR スキーマは、junii2 という日本の機関リポジトリのメタデータ交換フォーマットの代わりとなるものとして策定されました。junii2 は、図 1 にあるように、各大学の機関リポジトリから JAIRO（IRDB）にメタデータをハーベスティングする際に用いられているメタデータの交換フォーマットです。この JAIRO に収集されたデータは、CiNii や NDL Search、JaLC といったサービスにも流通しており、それぞれのサービ



（図 1）

スや DOI の登録などにメタデータが活用されるという流れになっています。

近年では、こうした国内のサービスだけではなく、OpenAIRE や BASE、OAIster といった海外のサービスにもデータが流通する機会が増えており、この junii2 は日本独自のメタデータのフォーマットのため、こうした国際的なフォーマットとの間にギャップが生じていました。

また、本日の SPARC Japan セミナーのテーマでもある、研究データといった新しいコンテンツへの対応を進めるために、今回、junii2 を改訂し、JPCOAR スキーマを策定するという流れになっています。

図 2 は JPCOAR スキーマの検討メンバーです。昨年度からタスクフォースを設置して検討を進めていました。国立情報学研究所の研究者と職員に加え、各大学の図書館員が一緒に検討する形となっています。昨年度は新しいメタデータフォーマットの方向性に関する検討や国際的な動向調査も行い、JPCOAR スキーマ案としてまとめたものをパブリックコメントとして公開しました。

今年度は、そのパブリックコメントで頂いたご意見の反映や、国際的な連携先との最終的な調整を経て、現在この JPCOAR スキーマの確定版を策定する作業を進めています。この確定版の JPCOAR スキーマは来月初旬に公開する予定です。

2.junii2 改訂の基本方針

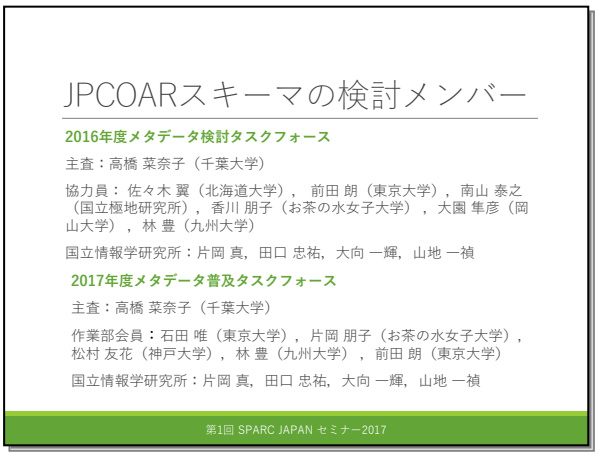
junii2 の改訂に当たり、主に三つの基本方針に沿って検討を進めてきました。1 点目は、近年、学術機関に求められているオープンサイエンス・オープンアクセス方針に対応した要素の拡充です。2 点目は、識別子の拡充・メタデータ構造の修正です。3 点目は、近年、国際的な学術情報流通の機会が増えており、日本独自の定義と国際的に主流な定義との間にギャップが生じていたため、国際的によく使われているメタデータの定義を採用し、相互運用性の向上を図ることです。

2-1.オープンサイエンス・オープンアクセス対応

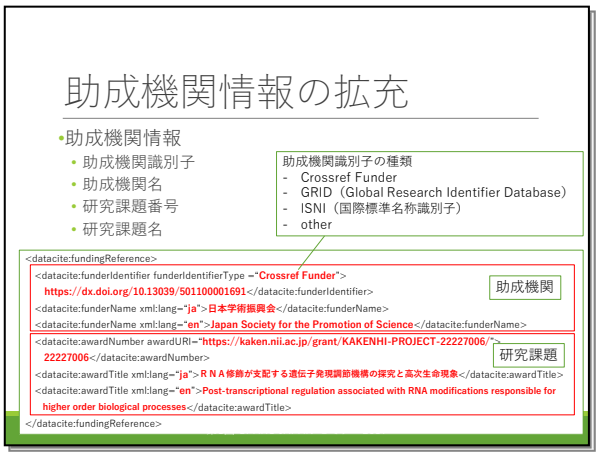
ここからは、それぞれの基本方針に対応して、具体的にどのような改訂を行ったのかご紹介していきます。

1 点目については、近年では政府によるオープンサイエンスに関する検討や、助成機関や大学によるオープンアクセスの方針の公開が進むなど、オープンサイエンス・オープンアクセスに対応するという社会的な要請が高まっています。このような要請に応えるために、学術成果をより公開しやすくなる要素や、研究データを管理するための要素の拡充を行いました。

追加したものは、研究助成を受けた場合の助成機関情報を記述する要素です（図 3）。内容としては、助成機関そのものに関するものと、研究課題に関するものとで構成されており、それぞれ助成機関名と研究課題名という文字列による情報だけではなく、助成機関の識別子、また研究課題番号を関連させて記述ができ



（図 2）



（図 3）

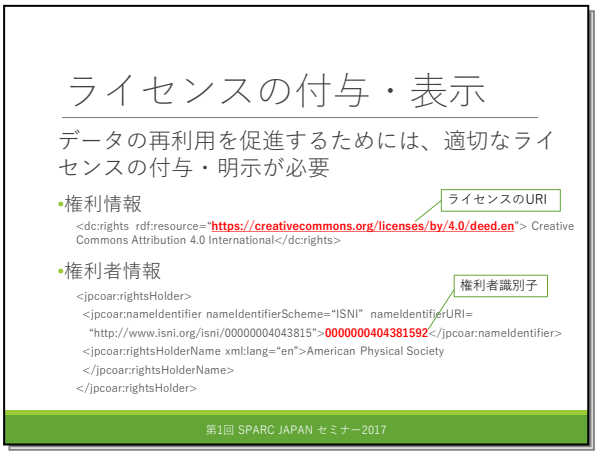
るようにしています。こうしておくことで、このデータが他のデータベースやサービスに渡った際にも、助成情報に関する分析や活用が進むようになると考えています。

もう一つのポイントとして、研究データを単純に公開するだけではなく、適切なライセンスと一緒に公開することで、その研究データの共有や再利用がより促進されるのではないかと思っています（図 4）。そのために権利情報と権利を持っている権利者に関する情報の二つを設定しており、それぞれ文字列で権利情報を書くだけではなく、権利情報の方は、例えばクリエイティブ・コモンズといった公開された標準的なライセンスも一緒に書けるようになっており、権利者の方は権利者の識別子を一緒に記述できるようにしています。

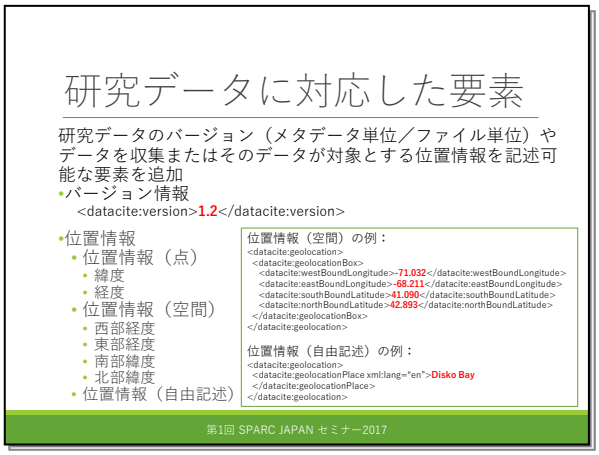
もう一つの側面としては、公的な研究助成を受けた学術成果は原則、公開を進めるようにという方向になってきているため、オープンアクセスの状態を確認できるアクセス権という要素と、論文処理費用（APC）の支払い状況を確認できる要素の二つを追加しました（図 5）。それぞれ、統制語彙と呼ばれる、決められた語彙で記述する形式で追加しています。

特に研究データは、細やかなデータのバージョン管理が求められる場合があるので、メタデータの単位と研究データの個別のファイル単位それぞれで、バージョンの情報を記述できるようにしました（図 6）。また、その研究データを採取した場所や研究データが示している内容の位置情報を経度や緯度を使って記すことも、文字列で書くこともできるようにしています。

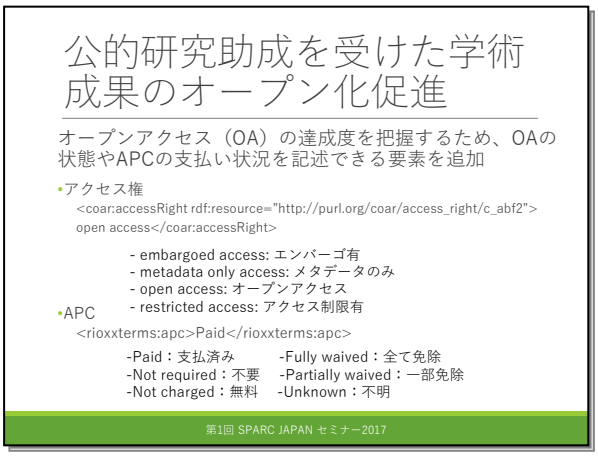
もう一つ、研究データに対応するために、寄与者の役割を属性として記述できるようにしました（図 7）。研究データの場合は、データの作成・収集など、論文の執筆よりも多くの人が関係してくる場合が想定され



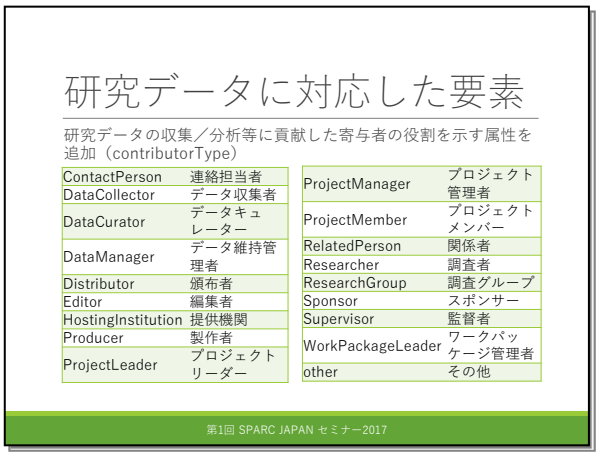
（図 4）



（図 6）



（図 5）



（図 7）

ます。そうした人たちの役割を属性として記述しておくことで、その貢献度を明示したり、適切な評価に結び付けたりするといった活用方法もできると思います。こちらも統制語彙の中から選択する形で記述するようになっています。

2-2.識別子の拡充・メタデータ構造の修正

改訂方針の 2 点目については、情報をより正確に記述できるように、識別子の拡充とメタデータ構造の修正を行いました。例えば、作成者と寄与者が複数名いて、それぞれが識別子や所属機関の情報を持っていたときに、それを並列に書いてしまうと、どの情報がどの人に結び付いているか、判別するのが難しいケースが生じてしまいます。そうした課題に対応するために、JPCOAR スキーマでは、関連する情報同士をグループ化して、階層関係を持たせて記述できるようにしました。

図 8 の例では、第 2 階層のレベルで人の情報を書いており、第 3 階層のレベルでその人が所属する機関の情報を書いています。こうしておくことで、この方が複数の機関に所属している場合も、第 3 階層だけを繰り返して記述するという書き方ができるようになっています。こうした構造化は他の要素でも必要に応じて行っています。

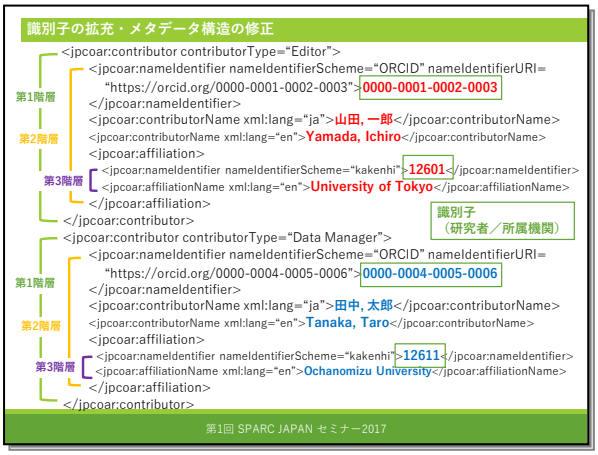
もう一つの改善事項としては、識別子と関連性に関する記述を増やしています。識別子というのは情報を一意に特定し得るものです。図 9 はメタデータ単位に付与するものの例です。メタデータ自体に付与された DOI、HDL、URI は資源識別子、外部のデータベースやバージョン違いの研究データに対して付与される識別子は関連識別子という 2 種類を用意しています。

また、研究データの場合は、バージョン違いの研究データにリンクさせたり、その研究データを用いて執筆された本や論文にリンクさせたりというように、関連に関する記述が増えることが予想されるため、関連識別子の関連性の種類も若干増やしています。

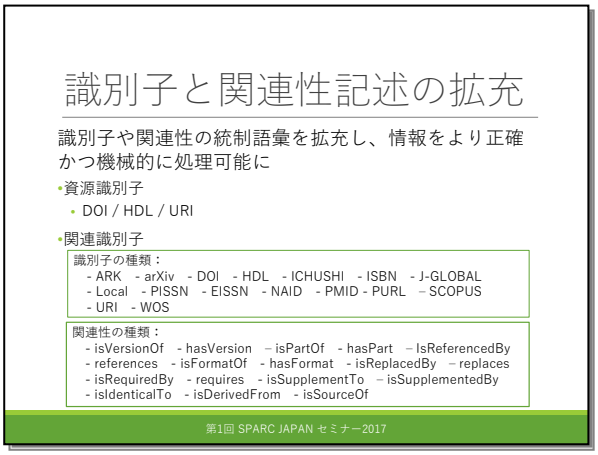
2-3.国際的に相互運用性の高いデータ交換のためのスキーマ定義

改訂方針の 3 点目に関連して、新しいメタデータフォーマットを策定するに当たり、国際的な動向調査を行いました。ただ、現在のところ、一つのメタデータフォーマットに準拠すれば全ての課題が解決できるという状況にはなかったため、JPCOAR スキーマは複数の海外などの定義を参考にした、複合的なメタデータ定義となっています。

具体的には、図 10 の 7 種類のスキーマ定義からそれぞれ要素を採用しており、研究データについては、研究データの国際的なスタンダードとなっている DataCite の要素を採用しています。その他にも、Dublin Core、COAR、OpenAIRE の統制語彙を採用しました。The Bibliographic Ontology からは雑誌の書誌記述に関する要素を採用しており、RIOXX からは APC の要素を採用しています。また、国内では、国



（図 8）



（図 9）

立国会図書館の DC-NDL の学位に関する要素を取り入れています。

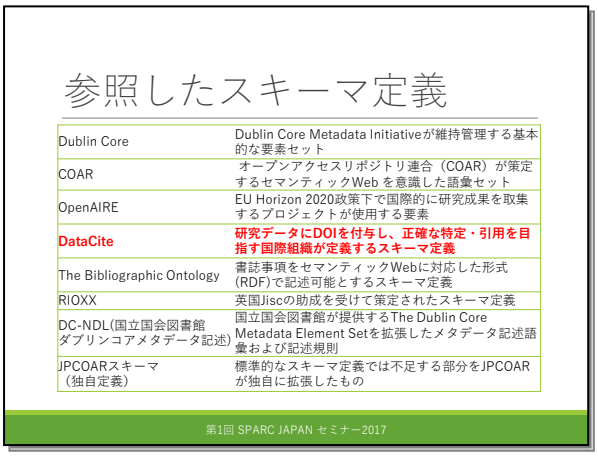
このように、できる限り外部の標準的な定義を取り入れる方向でまとめていますが、一部では、日本独自の事情や junii2 との互換性を保つために、どうしても独自に拡張せざるを得なかった要素も残っている状態です。

以上が JPCOAR スキーマの簡単な概要になります。

3.JPCOAR スキーマが果たす役割と可能性

ここからは、JPCOAR スキーマにどのような可能性があり、今後どのような役割を果たしていけるかというお話をします。

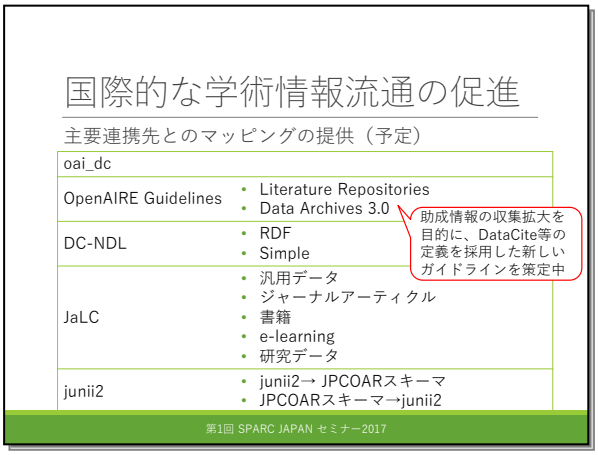
1 点目は、国際的な学術情報流通の促進です。現在の主要な連携先である OpenAIRE でも新しいガイドラインの改訂を行っており、お互いに情報共有をしながら改訂作業を進めています。かなり近い形にまとめられてきていますので、相互運用性は向上したと思います。また、研究データについては、DataCite という国際的なスタンダードから定義を採用しているため、もし今後、新たに研究データに関する連携先が増えた場合でも、ある程度の互換性は保てると思います。その他にも、図 11 のような主要な連携先とのマッピング表（対応表）をつくって公開する予定ですので、相互運用性の向上は図られていると思います。



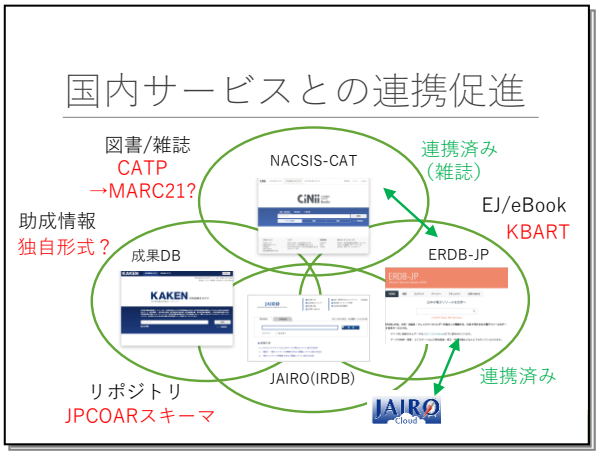
（図 10）

2 点目は、国内サービスとの連携促進です。JPCOAR スキーマは機関リポジトリ向けのフォーマットとなっており、そのプラットフォームである JAIRO についても、今年度末から来年度にかけて、JPCOAR スキーマに対応した改修が行われる予定です。研究データを扱うようになることで、こうした従来のリポジトリの連携先であった CiNii Articles などだけではなく、より新たなサービスが連携先の候補として拡大するのではないかと考えています。

例えば、図 12 のように、KAKEN のような成果系のデータベースには助成関係の情報や研究課題、その研究課題によって生み出された本や論文の情報が登録されていますし、NACSIS-CAT、CiNii Books には、研究データから生み出された成果物としての本の情報が登録されることも考えられます。ERDB-JP という、国内の電子リソースの情報を集めているサービスにも、研究データからつくられた学術論文が収録されている電子ジャーナルの情報が収録されるという可能性も考



（図 11）



（図 12）

えられます。このように、従来からこのようなサービスは割と近しい関係にあり、一部ではシステム連携も進んできていますが、そこに研究データが加わることによって、さらに関連性が高まってくるのではないかと思います。

ただ、それぞれのサービスで使われているメタデータのフォーマットは全く異なっており、連携を進めるに当たってはマッピングや対応表をつくる作業が必要になってくると思うので、今回の JPCOAR スキーマの策定の過程で得られた経験や課題を共有することによって、そうしたサービスの連携の促進につながればと考えています。

3 点目は、新しいサービスの展開です。研究データについては、データを管理をするだけではなく、それを広く一般に公開し、発見したりデータを再利用したりという仕組みが必要になってくるかと思います。図 13 は現在検討中のデザインのイメージ図を借りてきたもので、例えば従来の CiNii にはなかった「すべて」という検索メニューや、「研究」「研究データ」「人物」といったメニューが増えていることがお分かりいただけると思います。

「すべて」というメニューで検索した場合、キーワードに関連した研究データ、論文、本、人物、研究の課題といったあらゆる情報が一度に検索される予定です。従来の CiNii では、本は本、論文は論文という形で個別に検索するサービスでしたが、そこに研究データが加わることによって、本や論文やデータが関連性

（図 13）

をより高めてくると思いますので、このような網羅的な検索に対する需要も高まってくると予想しています。

「研究」というメニューで検索した場合、その研究に関連する研究データへのリンクや、関連するサービスである KAKEN へのリンク、またその研究によって生み出された文献・研究データの情報が掲載される予定です。さらに関連書籍、先行研究、被引用文献といった情報や、その他にも関連する論文、図書、研究というように、従来のサービスに比べて、かなり充実した内容の情報が掲載されます。

今後、研究データを扱うようになるなど、学術成果の情報量は増えていく傾向にある中で、関連する情報まできっちりサービスを提供することはなかなか難しくなってくると思いますが、例えばメタデータの中のある文字列からこうしたリンクを形成することよりも、メタデータの中に識別子や URI を含めておくことで、より正確な関連性の情報を提供できるのではないかと考えています。

「研究データ」で検索した場合、研究データへのリンクやクリエイティブ・コモンズといったライセンスへのリンクが表示される予定です。ライセンスの情報を文字列ではなく URI の形式で記述しておくことによって、サービスでもこのように正確なライセンスへのリンクが表示されるようになると思います。また、その研究データに基づく研究や論文、被引用文献、関連する情報も併せて表示される予定です。

「人物」で検索した場合、その人物が執筆した論文・本、関連する人物、関連する情報が表示される予定です。特に日本人の場合は、日本語や英語など複数の言語で著者名が表記される場合が多いと思いますので、メタデータにきちんとした識別子が含まれるかどうかによって正確な情報の集約が可能になると思います。

JPCOAR スキーマでは、国内的な普及状況を考えて、ORCID や ISNI といった国際的な識別子に加えて、日本で流通している e-Rad（科研費など日本の公的な助成事業で使われている研究者番号）や研究者リゾルバ

ーの ID も記述できるようにしています。

ただ今後、国際的な学術情報流通を考えた場合には、やはり ORCID のような国際的な識別子を付与しておくことの方がより望ましいため、今後の国内における識別子の普及状況にも注意を払う必要があると考えています。

4 点目は、研究データ管理の基盤としての波及効果です。junii2 は約 15 年間にわたって機関リポジトリのメタデータ交換フォーマットとして、かなり幅広く活用されてきました。現在のところ、約 645 の機関リポジトリが junii2 の形式に対応しているといわれており、JPCOAR スキーマはこの junii2 の基盤を継承する形となります。しばらくは並行運用の期間を設けますが、最終的には JPCOAR スキーマに移行していただく予定になっており、移行するときにまず、この機関リポジトリの管理者である図書館員がどうしたら JPCOAR スキーマに対応できるのかについて考える一つの機会になると思っています。

例えば、助成情報をどうしたら効率的に集められるのかといったことや、研究データにはライセンスを適切に付与した方が良い、メタデータに識別子・URI などを含んでおいた方が望ましいといった一つ一つの考え方の変更点が、研究データの管理や流通に関する意識の喚起にもつながればと思っております。

また、図書館員は研究データを持っているわけではないので、図書館員だけでは研究データ管理ができません。実際にこの研究データを管理している研究者の実情について理解を深めたり、研究者とコミュニケーションを取ったりする一つのきっかけにしていただければと考えています。

最後になりましたが、10 月 10 日に国立情報学研究所の会場をお借りして、JPCOAR スキーマの説明会を開催する予定です。ご興味のある方がいらっしゃいましたら、ぜひご参加をご検討ください。

●**フロア 1**　JST バイオサイエンスデータベースセンターの職員です。機関リポジトリ以外のサイトで JPCOAR のメタデータを付けて、それを例えば CiNii で検索できるようにすることは想定されているのでしょうか。

●**片岡**　現在のところは機関リポジトリがメインのターゲットになると思っていますが、やはり検討の過程で、例えばデジタルコレクションやデジタル化された図書についても、スキーマを考えていかなければいけないという課題は出ています。現段階ではリポジトリがメインですが、次の段階ではそうしたデジタルアーカイブや電子化された資料についても取り込めるようなことを考えています。

●**フロア 2**　専修大学の図書館員です。今日のお話はとても興味深いと思って聞いていたのですが、少し気になることがあります。今回、この junii2 を拡張した JPCOAR スキーマにいろいろ変更点が出てきていますが、それは機関リポジトリにデータを載せることを考えたから、junii2 の拡張のような形で考えられたのですか。

参考にされていた国際的な規格なども、データサイエンスに対応しているということでよろしいですか。「機関リポジトリにこういうものが出ました」という話になって、「データサイエンスやデータも載せられるようになりました」と言うと、やはり「載せてください」となってしまうので、そのときに、別のサーバー、システムの方で立ち上げた方がいいのか、それともリポジトリの中に論文とそのデータが混在しているような形を想定されていたのかを聞きたいです。

今回、CiNii の方が、Researh と Books と Articles が分かれているので、そういうことを考えたときに、各大学のリポジトリは共存してしまっていいのか気になったので、そのあたりを教えてください。

●**片岡**　一番参考にした OpenAIRE のガイドライン

でも、Literature Repositories という文献のリポジトリに、DataCite といった研究データの要素を追加して、研究データを載せられるようにするという流れがあったため、従来のリポジトリに研究データを載せるという方向性が一番自然なのではないかと考えて、リポジトリのスキーマである junii2 を拡張して、JPCOAR スキーマを策定しています。

CiNii の方はまだ検討中ということですが、資料種別ごとに個別に検索したい方は個別の検索をしていただけますし、網羅的に検索をするということも考えられているようです。これからの利用者のニーズがどのようになっていくかにもよるかと思いますが、私の個人的な感想としては、この「すべて」という検索がより使われるようになると期待しているところです。

●服部　SPARC Japan 事務局の服部です。CiNii の今の担当が出張に行っておりまして、前の前の担当としての発言をいたします。

CiNii Research では「すべて」だけではなく、論文（Articles）、本（Books）というくくりがまだ続くのかということですが、分けて利用したいというニーズがあるのだと思います。「論文を探したい」「本を探したい」ということもあると思うので、全て混ぜてしまうのはベストではないと思います。

まだ CiNii Research は検討の段階ですが、ここからは宣伝になります。11 月 7 日火曜日の図書館総合展の 10 時からのフォーラムでは「『次世代 CiNii』の展望」という企画を行います。9 月 20 日ぐらいから参加申し込みが可能になりますので、この話の続きが気になるという人は、ぜひ申し込みをお願いいたします。

●能勢　宣伝をどうもありがとうございました。

私から一つ伺ってよろしいですか。この JPCOAR スキーマは先ほどおっしゃったように元の junii2 を拡張したという形で、例えば研究者として、データをオープンにして登録したいと考えて、各大学の研究データリポジトリにデータを載せる場合、提供するデータも、JPCOAR スキーマに即したものを登録していくという形になっていくのですか。

●片岡　機関リポジトリが JPCOAR スキーマに対応している場合は、やはり JPCOAR スキーマに近い形で渡していただけると図書館員もそのまま取り込めて助かるところがあると思います。ただ、JPCOAR スキーマはどちらかというと流通用に割と項目を限定して、最小限の項目にしてつくったものですので、管理用にもっとリッチな情報が必要な機関もあるかもしれません。やはりそこは各機関でどういう項目でデータを集めるかを決めていただければと思います。

●能勢　そうなると、また先ほどの大澤先生の話にあった「いっぱい入力して、全部はねられて」という繰り返しになってしまうかもしれません。研究者側としては、やはり新しいスキーマなので、どういう情報が必要になってくるのかというのが知りたいところです。拡張したものは、助成機関の情報やオープンアクセスの状態、観測の位置などの情報が新たに必要になってきて、それを提供するということですね。

私は宇宙科学分野においてメタデータを自分で作成したことがあります。それはスペーススキーマという主に NASA でつくられたスキーマに従って書くのですが、この情報は必要でここに入力をしなければいけない、これは必須のマンダトリーで、これはオプショナルとか、やはりメタデータのスキーマを理解していくことが非常に大変で、もう二度とやりたくないのです。

それでいったん理解すると、今度はマイナーバージョンアップして、研究者としてデータをオープンにしたいのですが、とてもではないけれど、そうするためのメタデータを作成する気力が残っていませんでした。

先の話なので分からないかもしれませんが、実際に運用して JPCOAR スキーマでデータを流通していった場合、各大学においてどういう形で研究者がメタデータの情報を提供していくとうまくいきそうかという

見込みはありますか。例えば、Excel シートでフォーマットに入力していけばいいぐらいであれば、それほど大変ではないのですが、「xml を書け」「xml エディタをつくって研究者でやってください」と言われても、まずやりたくないですね。

●**片岡**　そうですね。このデータを研究者の方から頂きたいと思っているわけではなく、恐らく、お使いの機関リポジトリのシステムによって収集の仕方は変わると思います。例えば識別子だけ持っていれば、その他の情報は補完してくれるというサービスもあるかと思います。それはリポジトリのシステム側の機能として私も少し期待したいと思っている部分があります。その辺の情報交換も含めて、10 月 10 日に図書館員だけではなく、システムベンダーさんや研究者の方も含めた説明会と質疑応答する機会を設けておりますので、ぜひご参加いただければと思います。

●**能勢**　ご参加いただければと思います。各機関の研究データリポジトリが一般的になっていくと、junii2 の継承もあって、JPCOAR がメインストリームになっていくだろうと考えていてよろしいですか。この先どういった情報を提供すれば代わりにメタデータをつくっていただけるかという観点から質問をしているのですが。

●**片岡**　JAIRO で日本の機関リポジトリのデータを全て収集しているため、そのフォーマットが JPCOAR スキーマになるということで、今後は JPCOAR スキーマを目指して対応いただけるとありがたいです。

●**武田**　国立情報学研究所の武田です。個人的な立場から発言します。

今おっしゃった、このメタデータは誰がつくるのかというときに、システムベンダーを入れてという話ですが、そこは結構重要なところです。junii2 ぐらいのレベルであれば、今まで図書館員が自分で手で書いていた、書き写していたといったことが多かったのではないかと思うのです。junii2 はヒューマンリーダブルな世界でつくっていたので、構造もフラットでした。

しかし、この junii2 から JPCOAR になったら、もうかなり発想が変わっています。JPCOAR になった途端に、人間は普通ではほとんど読めない形式になるわけです。読むようにトレーニングすると読めるようになるのですが、普通には読めない。そこのところに大きな差があって、もうこれは機械処理を前提とするメタデータです。逆に言うと、そうだからいろいろ活用できるというメリットがあります。そういう意味では、つくる方も、人間が手でつくるというよりは、システム連携でデータを入れていくのが基本になるでしょう。そこで大きなステップに踏み込んだと私は理解しています。だから、必要な情報はシステムからうまく補完されます。ID を入れれば、もうその人の名前が補完されます。逆もあるかもしれません。

いろいろな情報を全て埋めないと駄目で、大澤さんが言ったように「全部フォーマットを埋めないと駄目です」と返されるということはなくなり、そこはシステムが頭良く働いてきちんと埋めてくれるようになるのです。

しかも、これは大学のファカルティ・マネジメント・システムと連携して、著者情報はそこから取ってくるといったシステム連携を前提として運用されることが期待されているのではないかと思います。最後の部分はそれをおっしゃったと思うのです。だからそのことを前提に皆さんがこれを使い倒していただけるようになるとうれしいなと思っています。

●**片岡**　言いたいことを詳しく説明していただいて、ありがとうございました。

●**逸村**　筑波大学の逸村です。さらにそれに乗るような感じなのですが、まず能勢先生が言われたマイナーバージョンアップについてです。これは研究者にとっては大変重要なのです。いわゆるアーリーアダプター

は、一生懸命こういうものが重要だと思って、研究者としていろいろつくります。しかし、それが知らないうちに変わっていたら、後で大変なショックを受けるのです。やはりそこは何らかのフォローが欲しいです。

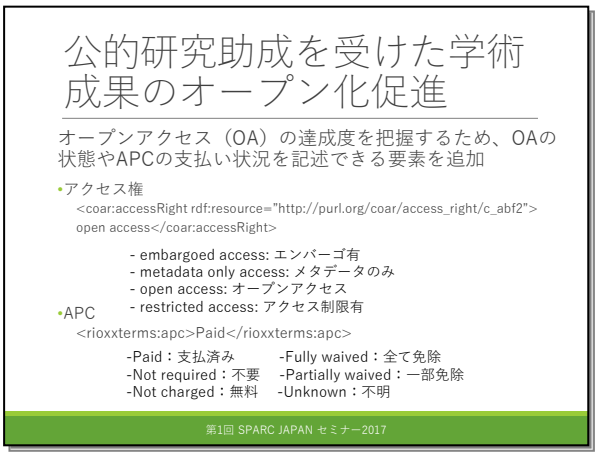
そのためには、今の武田先生のお話のとおりで、極端な話、ORCID を入れれば大抵のものは全部自動的にあちこちから集められるというぐらいにならないと、やはり研究者としてはちょっと付き合いきれません。特に若手は今、異動が激しいので、それごとにシステムが違うと「何だこれは」となると思います。老婆心でしょうか。でもやはり、そのためにぜひ頑張っていただきたいと思います。

●**片岡**　ありがとうございます。

●**能勢**　私が述べたマイナーバージョンアップについて、同じ思いを共有していただいて非常にありがたいです。バージョンが変わるとショックを受けて、どこが変わったかまた理解しなくてはいけません。二度とやりたくないですね。

もう一つ私からよろしいですか。細かい話なのですが、図 5 の APC というものがちょっとよく分からなくて、どういうものですか。データについて、paid や waived というのがあまりよく理解できませんでした。

●**片岡**　論文を投稿するときに、APC（Article Processing Charges）という論文処理費用を支払えばオープンになるというものです。



（図 5）

●**能勢**　ああ、一般的な APC ですね。そういう情報がなぜデータのオープンに必要になってくるか、あまり理解できなくて。別に APC は APC で、書誌情報でいいのではないか、データは関係ないのではないかと私は思ってしまったのですが、こういう情報も必要になってくるのですか。

●**片岡**　こちらもオープンアクセスがどの程度達成できているかを把握するにはどういう要素があればいいかを検討する中で出てきたものです。確かにメタデータにまで APC は要らないのではないかというご意見と、メタデータに APC も含めておいて、それを把握したり開示したりしたいというご意見の両方があり、今回は一応含めておく方向となっています。

●**能勢**　こういうメタデータの中身の説明は、Research Data Management（RDM）のマニュアルと何か連携はしているのですか。「この項目は RDM の資料の何ページに書かれている」などです。なぜかというと、「クリエイティブ・コモンズ（CC）」はここでは当然のように使っていますが、一般の研究者は、CC とは何か分からない、六つあるのも知らない、どれに該当するかも分からないという方も多いでしょうし、ISNI と言われても知らない方も多いと思います。

一般に広く研究者が使うには説明が必要です。メタデータを理解する、スキーマを理解するために、やはり何か資料が欲しいのです。APC に関しても、今私が思ったように「これは何だろう」と思いますので、そういうマニュアルや簡単な説明をつくられる予定などはございますか。

●**片岡**　一応、今、ガイドラインというものを作成しているのですが、それは主にリポジトリを管理する図書館員向けの内容になっています。研究者の方にこんなに関心を持って見ていただけるとあまり思っておら

ず、申し訳ありません。

●**能勢**　この部屋の中でだけかもしれませんが。

●**片岡**　研究者向けの教材との連携はまだできていませんが、確かにそういうものもあると良いかと思いましたので、今後の検討の課題とさせていただきます。

●**能勢**　この部屋にいる研究者は「メタデータのスキーマ」と言っても興味を持つのかもしれません。

●**片岡**　ありがとうございます。

第 1 回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係：研究データの管理と流通から考える」

研究データ利活用協議会（RDUF）紹介

武田 英明

（国立情報学研究所/研究データ利活用協議会）

講演要旨

研究データ利活用協議会（RDUF）は国内の研究データの共有や公開に関わる関係者の集まる場として、ジャパンリンクセンター（JaLC）を母体に発足したものである。ここでは、本協議会の活動の紹介と新たな取り組みである小委員会制度の説明を行う。小委員会は個別トピックスの議論の場として用意したもので、現在、企画を募集中である。

武田 英明

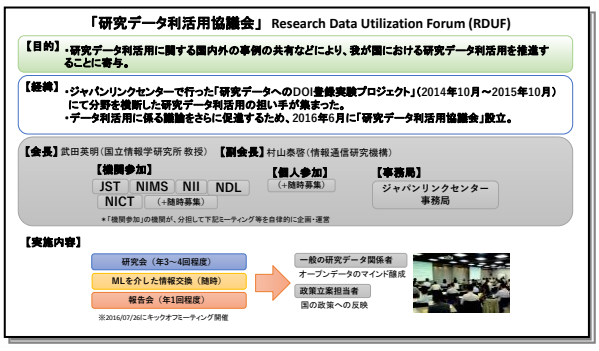
http://www.nii.ac.jp/faculty/informatics/takeda_hideaki/

今回の SPARC Japan セミナーは、研究データ利活用協議会（RDUF）が共催させていただいているので、RDUF の PR の時間を設けていただきました。RDUF の会長という立場でお話しします。

研究データ利活用協議会（RDUF）の概要

RDUF は、研究データ利活用に関する国内外の事例の共有などにより、わが国における研究データ利活用を推進することに寄与することを目的とした任意の組織です（図 1）。

設立のきっかけは、ジャパンリンクセンターが行った、研究データに DOI を付与する実験プロジェクトです。その前に、DOI の仕組みについて説明します。DOI の胴元は International DOI Foundation（IDF）ですが、IDF は DOI のデータベースの管理のみ所管して、それ以外は IDF から Registration Agency（RA）として許可を受けた組織が、実際に DOI の登録などの作業をしています。その最大手が CrossRef です。IDF の立場から見ると CrossRef も一つの RA で、ジャパンリンクセンターも一つの RA です。ジャパンリンクセンターが 2012 年に設立されて、国内で DOI を付与する業務を行っています。



（図 1）

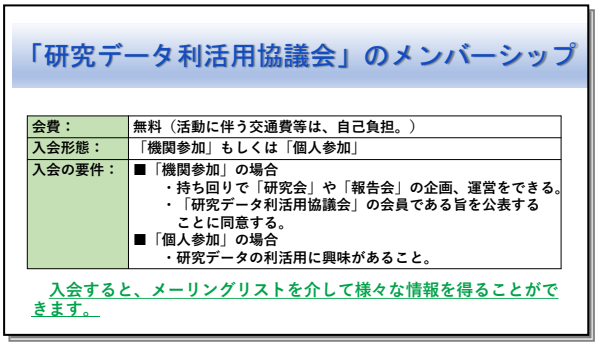
ジャパンリンクセンターとは、国立情報学研究所（NII）、科学技術振興機構（JST）、国立研究開発法人物質・材料研究機構（NIMS）、国立国会図書館（NDL）という四つの機関が共同で運営する形で設立されたものですが、実際の事務局およびシステムの運用等は JST が担っているという仕組みで運営されている組織です。正確にいうと、ジャパンリンクセンターは独立した組織ではなく、あくまでもプロジェクトという立ち位置です。ちなみに CrossRef は NPO 団体です。片岡さんの JPCOAR スキーマの話で出てきた DataCite も、DOI の RA の一つです。

ジャパンリンクセンターは、今までは主に国内の論文に DOI を付与してきましたが、公開したい研究データに DOI を付ける、「研究データへの DOI 登録実験プロジェクト」を 2014 年 10 月から行い、データに DOI を付けたい人は一体どこにいるのか、そもそも公開するデータを持っている機関はどこなのかということを手探りで調べました。そうすると、幾つかの機関が集まりましたが、このプロジェクトは 1 年間で終わりました。もう少しこのコミュニティを発展させたいということで 2016 年 6 月に始まったのがこの研究データ利活用協議会（RDUF）です。「協議会」という名前が付いていますが、これはジャパンリンクセンターの特別部会で、ジャパンリンクセンターの一活動という位置付けです。

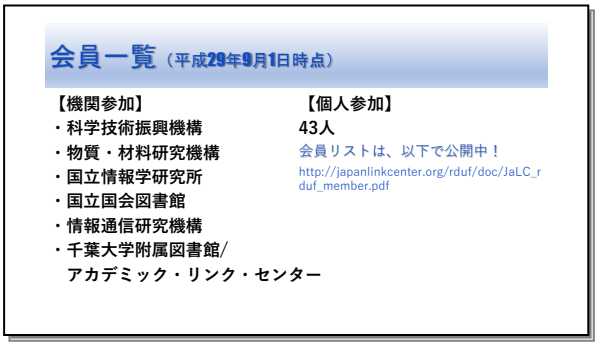
ただ、本体のジャパンリンクセンターとの関連はそれほど深くなく、むしろオープンにいろいろな人に入ってもらいたいと考えています。現在は機関参加として、ジャパンリンクセンターの共同運営機関である NII、JST、NIMS、NDL に加えて、情報通信研究機構（NICT）、千葉大学のアカデミック・リンク・センターに参加いただいており、産業技術総合研究所（AIST）にも参加いただく話を進めつつある状況です。あとは個人参加で興味を持つ方は誰でも入れます。

RDUF は設立から 1 年少したちました。実施内容は、勉強会や研究会、メーリングリストを介した意見交換、報告会です。参加者は、研究者はもちろん、研究機関の担当者、大学の図書館等の関係者といった研究データ公開に関わる多様なステークホルダーを想定しています。

入会形態は、機関参加もしくは個人参加で、入りたい人は入ってくださいという形ですが、機関参加はむしろ企画側に回り、研究会などを企画するという立場です（図 2）。個人参加者はメーリングリストに入ります（図 3）。



（図 2）



（図 3）

（図 4）

RDUF の活動実績

RDUF は今まで図 4 のような活動を行ってきました。第 2 回研究会として、昨年度も SPARC Japan セミナーを共催しました。大きな活動としては、Research Data Alliance（RDA）の参加報告会があります。RDA のような世界の活動への日本からの参加をうまく盛り上げたいというのが、設立のもう一つの動機でした。公開シンポジウムも行いました。

今年は 11 月 6 日に、国会図書館に企画していただき、場所もお借りして第 1 回の研究会を行います（図 5）。これは RDA の第 10 回報告会や iPRES 報告会なども兼ねていて、みんなで情報共有しようという企画です。

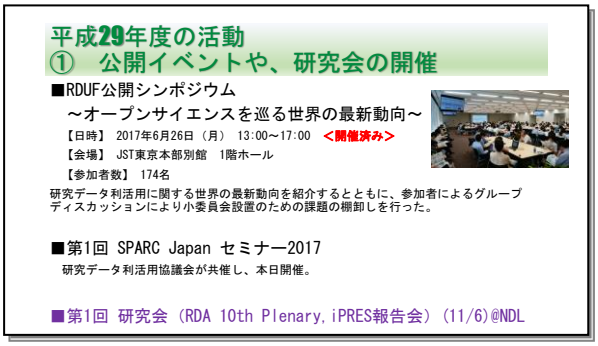
小委員会の設置

また、今年から小委員会をつくってみようと考えました。全体の研究会以外に、有志で集まった小委員会で、あるテーマにもっとフォーカスして、1 年ぐらいかけて議論するのです（図 6）。

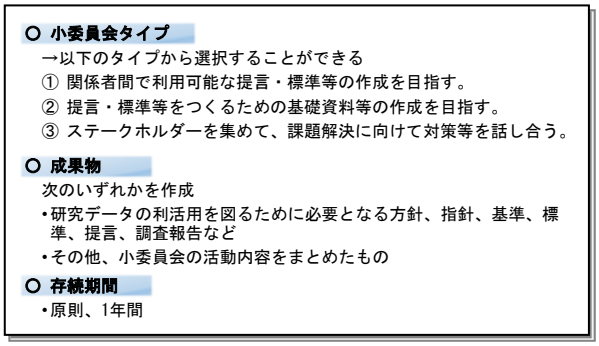
標準をつくるという高い目標を掲げていただいてもいいし、単に提言あるいは基礎資料をまずつくろうということでも構いませんが、とにかく 1 年ぐらいかけて成果を出してもらいたいと考えています（図 7）。

提案書のフォーマットもつくり、検討テーマを 9 月 29 日まで募集しているので、やりたいという意思を表示していただきたいと思います（図 8）。今のところ「研究データ・マネジメント・プランについて一緒に考えましょう」という提案が既に出ています。

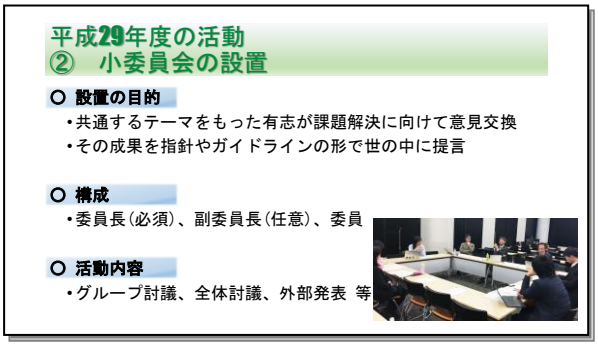
オープンアクセスリポジトリ推進協会（JPCOAR）にも小委員会のようなところがあるのですが、それはやや会員制で、図書館系のメンバーが比較的多いです。それに対して、RDUF の小委員会はオープンなので、もっと広いオーディエンスで、あるテーマに絞って議論したいのであれば、この小委員会制度を使っていただければと思います。JPCOAR の活動とオーバーラップさせても構いませんし、SPARC Japan の活動とオーバーラップさせても構いません。何せ任意の組織なので、そのあたりは自由に考えます。



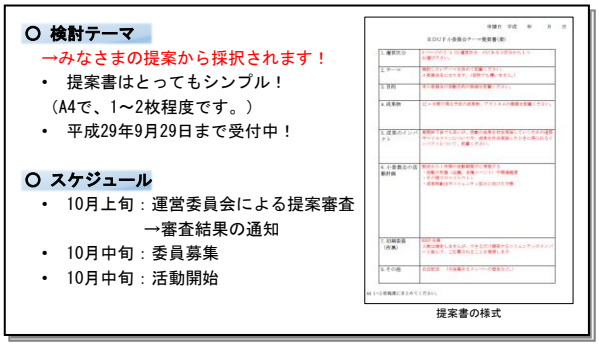
（図 5）



（図 7）



（図 6）



（図 8）

「委員会」という堅苦しい名前を付けたのは、その方が所属組織などに説明しやすいだろうというぐらいの意味で付けたもので、実際は自由に活動してもらいたいと思っています。もっと深い議論をしたい方には、ぜひこの協議会を活用していただければ幸いです（図9)。

（図 9）

第1回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係：研究データの管理と流通から考える」

全 体 議 論

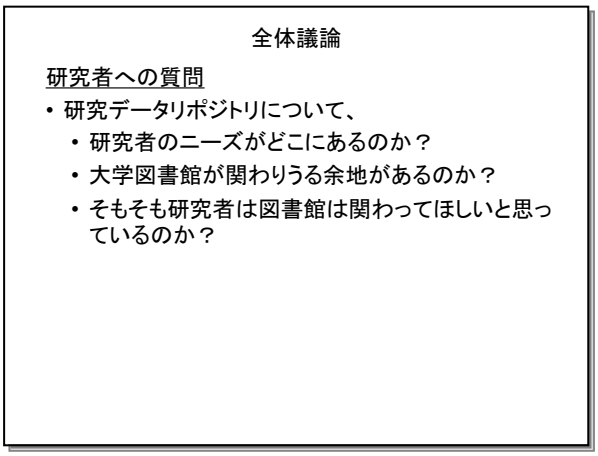
能勢　正仁　（京都大学大学院理学研究科）
倉田　敬子　（慶應義塾大学文学部）
大澤　剛士　（国立研究開発法人 農業・食品産業技術総合研究機構 農業環境変動研究センター）
西薗　由依　（鹿児島大学/JPCOAR 研究データタスクフォース）
片岡　朋子　（お茶の水女子大学/JPCOAR メタデータ普及タスクフォース）
武田　英明　（国立情報学研究所/研究データ利活用協議会）

●**能勢**　ここから全体議論に移ります。Twitter でご参加されている方で、もし何か質問がある場合は受け付けますので、つぶやいてください。

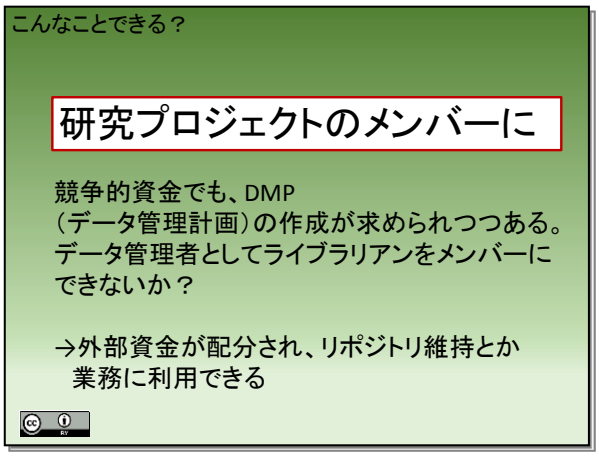
図 1 は、最初に研究者への宿題という形でお願いしたテーマです。これについて聞く前に、5 人の先生方に登壇してお話しいただきましたが、これについてもう少し話を伺いたいことや質問はございませんか。

●**服部**　SPARC Japan 事務局の服部です。「大学図書館が関わり得る余地があるか」について、大澤先生の資料にも、図書館員が研究プロジェクトに関わるといったものがありましたが（図 2）、こういうものを組織として受けるのはまだ難しいと思うのです。どのぐらい業務量があるか分からない、図書館員は 2〜3 年で別のところに異動するので、ちゃんと続けられるかどうか自信がないところがあると思うからです。そこで先生方にお尋ねしたいのが、個人で受けて異動したら「ごめんなさい」でも許されるのかどうかです。ご意見いただけますでしょうか。

●**大澤**　まずその話は、リポジトリ、箱がどこにあるかということさえ担保されれば、異動はもう普通のことだと思います。若手・中堅ぐらいの研究者は異動するのが当たり前ですし、一つのプロジェクトの途中に 2〜3 回異動するなんてことも当たり前ですが、そのテーマ自体は抱えていくわけです。例えばプロジェク



（図 1）



（図 2）

トに組織として関わるとすれば、どうなるか何とも言いにくいところがありますが、仮に個人レベルで関わるとしたら、それ自体はそんなに問題ではないという気がします。

ただし、所属する組織が替わったとき、研究者の場合は研究費を背負って来るので、研究者として異動する限りはそのプロジェクトをやめろということはないでしょうけれど、図書館員の場合はそうはいかない可能性が想像としてはあると思います。

●**服部**　例えば同じ大学の図書館の隣の係などでしたら、少しやれる余地があると思うのですが、特に国立大学では、隣の隣の県など、かなり大規模な異動があるので、そうすると仕事を続けるのが難しいと思います。そういうところで思い悩むよりも、先生方とぶつかって思い切ってやったほうがいいのかなと。

●**大澤**　今のはちょっと私の答えがずれていたと思ったのですが、例えばわれわれは農林水産省の委託プロジェクトのようなものによく関わります。その場合、われわれのような国立系のところは、所内での人事異動的な意味の異動はあまりないのですが、都道府県だと、次は普及所にいた行政の方が研究の担当になるということがよくあります。そのプロジェクトの期間内はずっとその県が参画しているけれど、担当が入れ替わることは比較的頻繁にあるということです。それで全てすんなりいっているとは言いませんが、比較的それでもとにかくプロジェクト自体は運営されるので、何とかなるのではないかという気はしています。

●**服部**　ありがとうございました。

●**能勢**　私は倉田先生に伺いたいのですが、今日ご紹介されたインタビューの結果は、少し年配の方を対象にされた結果だとおっしゃっていました。もう少し下の若い世代はどう考えているか、何かコメントはございますか。もう少し若い人だと違うとか。

●**倉田**　インタビュー対象者は、もちろん全員がいわゆるシニアではなく、助教クラスのお若い方にもお聞きしています。そのときに感じたのは、やはり若い方のほうが大変というか、少し余裕がないということです。とてもそこまで思いが至らないのです。

お年を召したと言うのはちょっと失礼ですが、そういう方は、「自分がこれまで持っていたデータは、定年になったらどうなるのだろう」とすごく心配なさっています。自分の研究がどうこうというよりも、「このデータは絶対にその後使えると分かっている。けれども、自分の後進が、この大学の同じ研究室に教授として残ってやっていけるという保証はない。そうすると、このデータをどこに託せばいいのか、すごく心配だ」と言っていたことが大変印象に残っています。それと似たような話は他でも聞いたことがありますし、私自身も少しそう思います。

データリポジトリは、自然科学系では幾つか出てきていますし、社会科学系も東大のものなど幾つかありますが、全国の、誰でもすぐにデータをぱっと出せるというものはないのです。そうすると、自分が「これは結構使えるのではないか」と思っているデータでも、欧米の figshare や Dryad などに入れるしかなく、それでいいのかというのは、やはりすごく思うところがあります。

年齢は関係なく、データを蓄積して持つようになると、これはどうしたらいいのかという心配が出てくると思います。ただ、例えば今、私が「慶應の機関リポジトリにこのデータを入れますか」と言われたら、困ってしまうと思います。メディアセンターに「これ何とかして」と言っても、どうしようもないという話なので。

先ほど、図書館員の業務としてはなかなか難しいとおっしゃったのですが、そうなると逆に、すごく親切な図書館員に雑用を助けてもらうということしかできなくなって、それは少し違うだろうと思うので、私は業務としてきちんと道をつくってあげない限り、できないのではないかと思います。

●**武田**　今のお話は、シニアにはシニアの悩みがあって、若い人には若い人の悩みがあるという話で非常に面白かったです。シニアには、確かにリタイアしたときにデータはどうなるのかという問題があると思います。死ぬまで自分だけで抱きかかえたい人は別として、データを残していきたいという人たちに手だてを提供してあげることも、機関リポジトリなり、リポジトリのある種の役割だと思います。

リタイアすると時間ができるので、きちんとキュレートできるような気もしますが、大学の機関リポジトリに「元教授はもう入れられません」と言われてしまうと困ります。だから、そこの感覚も「元スタッフだったら入れてもいいですよ」となると、これからは65 歳過ぎたぐらいの方が頭も元気だし、時間もあると思うので、そういう方のエネルギーをうまく使ってデータの整備や公開もできるといいなと今の話を聞いていて思いました。

ちなみに今、多くの大学ではそういう運用はなされていないのではないかと思うのですが、どうですかね。

●**能勢**　そういう前例など、ご存じの方はいらっしゃいますか。

今、倉田先生と武田先生がおっしゃった話は、私の自然科学、宇宙科学の分野でも話題になったことがあります。税金を使って取ったデータで、貴重で、まだ使えるのに、その方が定年されると、公開したり保存したりする術がない。そのときは、学会で取ったデータなので、学会でそういう仕組みができないかという話になったのですが、それをするにも費用がかかるので難しい。そうなると、大学の機関データリポジトリに登録できるといいのですが、今おっしゃったように、確かに定年されるとそこまで責任が持てないという話も出てくるかもしれません。でも、確かに定年でいらっしゃらなくなる先生のデータをどのように保管するのかというのは大切な話ではあります。

もう一つ、若い方のほうが余裕がないというのはあり得ることだと思いました。大澤先生もおっしゃっていたように、本来の研究のほうで手いっぱいでそこまで手が回らないということです。先ほど、「データを管理して公開していくのは当然だ」という形で、若い人へのリサーチデータ・マネジメントの教育に使えるという話がありました。これは非常に大切な視点だと思いました。

●**林賢紀**　国際農林水産業研究センターの林です。今の定年した方のデータの話で思い出したのですが、九州大学に、1970 年ぐらいの博士論文のデジタル化ができないかとこちらからお願いして、電子化して向こうのリポジトリに入れていただいたという経験があります。

九州大学の OB から、「自分の持っている博士論文を電子化したいが、どうすればいいか。お金は幾らでも出す」と聞かれたので、「まず九州大学さんに相談してみますよ」と言って相談したら、九州大学に「では、うちの成果だから、うちで電子化します」と言っていただけました。判型も大きくて少し大変だったと聞きましたが、ご本人から電子化してかまわないと言っていただけるなら、そこはさらっといけるかと思います。確かに、研究データを抱えていて、「これ、死ぬ前に何とかしたいんですけど」というのはある話なのではないかと思って聞いていました。

倉田先生のお話のとおり、たまたま親切な図書館員がいてやってくれるのと、組織として、「電子化はわれわれのミッションだからやるのだ」という形でやるのだと、できれば後者のほうが、お願いするほうもすっきりできるのでいいなと思いました。

●**能勢**　今の定年された先生のデータで思い出したのですが、「ダークアーカイブ」という言葉を聞いたことがあります。京都大学でもダークアーカイブ用の、もう二度と上書きできない、固定したデータをずっとアーカイブしていくという試験的なものがあり、そういうものを利用できると、図書館でも定年された先生のデータを保管していけるかもしれないと思いました。

●**林和弘**　NISTEP 林です。今の議論は、定年後のデータに限らず、若い人が研究対象を変える際にも起きる議論だと思っています。特に化学の世界で、私が知る限り、一つのテーマで研究できる期間がどんどん短くなっています。「ガリウムをやっています」と言えば 40 年研究できていた時代が私の大先輩のころありましたが、今は 5〜10 年たったら、テーマを大きく変えなければいけません。

北海道大学のオープンサイエンスのシンポジウムでもお話ししたのですが、研究者は次の研究対象が見つかると、できるだけそちらにリソースを注力したいという性質があるので、定年後というよりは終わった研究のデータを、「またいつでも再開できるように置いておけます」「安心して管理できます」というものができると、Win-Win で、大澤さんが言うところの手をつなげる状態ができるのではないかと思います。そう言うと、大体、研究者側からも図書館側からも「それだったらいけるかもしれませんね」と言われます。

これであれば、研究者は新しいことに専念できますし、図書館は今動いている研究に対しては自分がコミットできるかどうか非常に抵抗感がありますが、終わったものであれば、終わった範囲のところで最低限のメタデータを付けて、ラッピングしてどこかに置いておくことができるなど、戦略が取れてくるのではないかという話になったことがあります。

今の倉田先生のお話をもう少し一般化して、研究が終了して次のテーマに移るときに、研究者と図書館の方が手を握れることがあるのではないかということで、指摘させていただきました。

●**能勢**　そうですね。定年退職してしまうと、所属員ではないという問題が、武田先生のご指摘のようにあるかもしれませんが、いったん終わった研究の変わらないデータについて、積極的に図書館の方々にご相談してアーカイブしていただくのは、一つ今後の方向かもしれないと思いました。

他にございますか。この話に限らず、他のご質問でも結構ですが。

●**Twitter 担当**　Twitter での質問です。「研究プロジェクトに図書館員が加わるという話ですが、実際のところ、どうやってプロジェクトに必要な図書館員を見つければいいのでしょうか」。

●**倉田**　見つける見つけないというのは、あまり関係ないのではないかと思います。そういう役割が本当に認められているのであれば、逆に言えば図書館員は誰でもやれなければいけないのではないかと思います。もちろん、図書館員の中にいろいろな役をやる人がいるという体制になっていれば、その中である役割を担っている人、今で言えば、例えば Reference Librarian のような業務・サービスをやっているところに依頼が行くというイメージです。

私が思っている Embedded Librarian のようなものは全部そうです。あるサブジェクトがあったら、そのサブジェクトを図書館が受けて、誰を派遣するかはその後の話で、誰かを一本釣りにして「この人が欲しい」というイメージではあまりありません。図書館のサービスとしてやるからには、誰をどうやって見つければいいかではなくて、そういうサービスがあって、ただ申し込めば派遣してもらえるというのが、本来のサービスの在り方だと思います。

●**大澤**　私の考えとしては、研究者的な発想なので図書館員の方には受け入れにくいかもしれませんが、研究者だったら誰とでも共同研究できるかというと、当然違うわけです。

それと同じように、図書館員だったら誰でもできるというわけではないと思うので、よその大学間で組めるかということは置いておいて、今日のような機会や学会で、「お互いが持っていないスキルを持っているから一緒にやりましょう」「同じような興味を持っているから、相乗効果を狙って一緒にやりましょう」「あの人はこういう技術を持っているから、力を借り

たい。一緒にやりましょう」「これだけのデータリポジトリを進めているということは、あそこの大学は目利きの人がいるに違いない。ぜひ一緒にやりませんか」という話になるのではないかと考えていました。ただ、それは非常に研究者的な発想なので、一般的に実現可能かどうかは分かりかねます。

●**能勢**　今の質問に対して、研究者の方でコメントはございますか。

●**逸村**　筑波大学の逸村です。当然、気心が知れている、この人となら一緒にやれるというのは大変重要なので、大学図書館の業務としていただくなら、管理職の方はそういうマッチングを重々考えてほしいです。また、一本釣りもないわけではないですが、そうするとその部署でのその人の立場というものもまた別途あるでしょうから、結構微妙なところです。やはりこれは管理職の仕事だと思います。

先ほどの服部さんの質問に戻ります。機関リポジトリの担当者が 2〜3 年で動くのはやむを得ないのですが、単にデータを預けるだけではない話を研究者としても持ち込みたいことがあります。その際にも組織として「あの人がいなくなったから、もうこの質問は受け付けません」というのは、何とかしてほしいというのは実感としてあります。

●**能勢**　今のお話に関連したコメントなどはありますか。管理職の方、いらっしゃいますでしょうか。

この企画を立てる段階で図書館の方と話をしていると、やはり共通の疑問がありました。それが「研究者のニーズがどこにあるのか？」「大学図書館が関わり得る余地があるのか？」「そもそも研究者は図書館に関わってほしいと思っているのか？」です。今日のご講演でかなりの部分は答えが出たかもしれませんが、研究者、これに対して、私見でも結構ですので、お答えいただけるとありがたいです。

●**深貝**　横浜国立大学の深貝です。研究者のニーズは、率直に言うと、自分のテーマに都合がいいように図書館が助けてくれるという非常にわがままな話になってしまいます。図書館という機能には、サポートできることとできないことがあるわけです。それも親しい人だったらサポートして、そうでなければサポートしないというえり好みを図書館側がするわけにはもちろんいきません。そうすると、広く薄くみんなにできるサポートはあるのかというと、よく分からないのです。

ルーティンワークで図書館が動いていくようになればいいかというと、現代の学術や大学機能の変化、そして何より情報化によって、図書館の担う役割が急激に変化しているので、今までどおりの古い図書館でいてもらっては困ります。そして、研究者もやはり変化の中でうろたえているのです。うろたえている人々は助けてほしいと思うけれども、自分の都合でしかうろたえていないから、非常にローカルな注文を付けてしまいます。ローカルな注文に満遍なく図書館が応えられるわけがありません。結局、非常にフォーカスされたトピックスでサンプルをつくるというのは、一つの参考にはなるとは思いますが、それをあらゆるところでできるわけはないのです。

今、データの管理の仕方については、機関リポジトリをベースにそろえましょうとなっているのは、日本が世界的にも突出して機関リポジトリを既に持っているからです。この機関リポジトリが使えるということで、文部科学省のオープンサイエンスの方針ができる 1 年前、内閣府の報告書が上がったところで、「機関リポジトリを頼りにオープンサイエンスを進めていきましょう」と言ったから、いつの間にかオープンサイエンスが図書館の仕事になってしまったわけです。

でも、図書館がそれをいくらやっても、研究者がオープンサイエンスとは何かということを理解して、かつ知識をオープンにしたがっているかというと、どうもそうはなっていません。しっちゃかめっちゃかなわけです。この中で、研究者の期待をそのままともに受けて、「旧式のやり方の都合のいいことをやってく

れ」と言う人に応える必要があるかというと、そうすると逆に図書館の変化が遅くなってしまいます。今後のことを見通した上で、図書館が次のステップにつながるようなスキームをどうつくっているかというメッセージを出していく方が先になるのではないかと思います。

ところが、これは大学によってまちまちですが、さまざまな人的配置の中で、あるいは大学の将来的な運営の見通しの中で、スタッフの数をどんどん減らさざるを得ないようなところがあります。少ない人数で図書館スタッフが切り盛りせざるを得ないとなると、一つ一つのローカルなユニットから成る図書館で、あらゆる機能の研究者の注文を受けられるかというと、そうではありません。そうすると結局、複数の図書館横断的なスキームができるようなストラテジーを考えるということになります。機関リポジトリの横断組織にもなるようなオープンアクセスリポジトリ推進協会（JPCOAR）でできるかというと、これはやはりデータを集められるうまい仕組みをつくろうということであって、サービスの個々のものを JPCOAR に持ち込んでやってもらうというわけには到底いかないわけです。

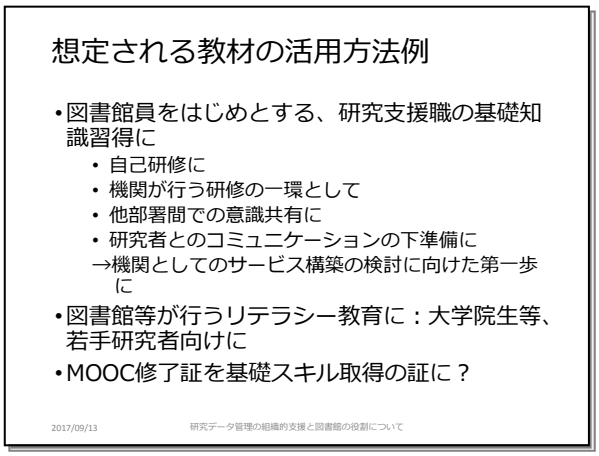
結局、その間になる中間的なプラットフォームやサービスのモデルなどをどうつくっていけるか、そして、活きのいい図書館スタッフ、若手の研究者が新しい試みができるようなチャンスをどうやって開いていけるかということを図書館が比較的柔軟に考えられるようになるとよいと思います。

手書きと謄写版印刷のカタログではない、電子的 OPAC をつくったような感覚がある人が、今の新しい変化の中で、どういうものをつくらなければいけないかをサジェスチョンして、変化の中で図書館が動いていくべき方向を探る、これがしばらくの課題ではないかと思います。

●**能勢**　確かにおっしゃるように、広く浅い誰にでも適用するようなサービスは、すぐにはできないと思います。やはり実例を何か一度やってみたいというのは、私がずっと希望していることで、JPCOAR の教材の第 8 章に実例がもし加われば、非常にありがたいです。実例をお願いするにしても、図書館員の方がそういうことをウエルカムなのかというのは、研究者側としてはいつも気にしていたことです。それは多分、大澤先生も同じで、図書館の方にとってインセンティブがあるのかということになるわけです。

今日は実際に生の声を聞かせていただいて、非常に勇気付けられた言葉が、西薗さんからの「図書館員にとって、利用者の役に立つことは無上の喜び」だったと思います。また、JPCOAR の教材の活用方法例として「研究者のコミュニケーションの下準備に」ということも挙げられていたので（図 3）、図書館員の方としても受け入れてくださる余地は既にあるのではないかと考えましたが、すぐに全体に通じるようなシステムや仕組みはなかなかつくれないと思うので、このような場で、興味を持っている研究者と図書館員の方とで、何か具体例を一つつくってみたいというのが、今回改めて思ったことです。

この全体議論で、研究者への質問に対して研究者が回答したかどうか分かりませんが、私としては、企画をした人間が一番得をしたような気がしています。専修大学や京都大学といった図書館の方の生の声も聞かせていただきましたし、図書館員の方の研究データのオープン化に関わるときの心構えは、非常にオープンマインデッドで、「利用をしていただけるならウエル



（図 3）

カム」だと捉えたのですが、間違っていたら訂正してください。研究者として、データをオープンにしていきたいという者にとっては、非常に勇気付けられました。ありがとうございました。

SPARC Japan NewsLetter
ISSN 1883-826X
NO.34 2018 年 3 月

SPARC Japan ニュースレターでは、各回セミナーの報告に講演や全体議論を書き起こしたドキュメントを加え、さらにそのほかの SPARC Japan の活動をご紹介します。

CONTENTS

■ SPARC Japan 活動報告
SPARC Japan 運営委員会
SCOAP³ 支援

■ SPARC Japan セミナー報告
概要
参加者から
企画後記
ドキュメント
（講演・全体議論）

■ SPARC Japan 活動報告

CC BY

SPARC Japan 運営委員会

SPARC Japan 運営委員会の会議資料をウェブサイトで公開しています。
http://www.nii.ac.jp/sparc/about/committee/

SCOAP³ 支援

第 5 期 SPARC 事業の達成目標の一つとして「国際的な OA イニシアティブとの協調」および「学術情報流通にかかわるアドボカシー活動」を掲げ、その具体的活動として SCOAP³（スコープ・スリー）を支援しており、日本参加機関の取りまとめと広報活動を行っています。

SCOAP³（Sponsoring Consortium for Open Access Publishing in Particle Physics）は、スイスの CERN（欧州原子核研究機構）が中心となり高エネルギー物理学分野の査読付きジャーナル論文のオープンアクセス化を実現することを目的とした国際連携プロジェクトで、日本では高エネルギー加速器研究機構、国公私立大学図書館協力委員会、および国立情報学研究所（NII）が 2011 年に共同で関心表明 EoI に署名し、NII が National Contact Point となりました。

フェーズ 1（2014 年～2016 年）における日本の参加機関は 34 機関で、フェーズ 2（2017 年～2019 年）は 2017 年に 40 機関、さらに 2018 年からアメリカ物理学会が対象誌として加わったことをうけ、67 機関が参加を表明しています。「大学図書館と国立情報学研究所との連携・協力推進会議」のもとに設置されるタスクフォースでは、拠出金額の算定方法を定めるほか、参加意向と拠出金額の確認などを行うとともに、物理学研究者向けにも広報活動を行っています。また日本参加機関の拠出金は毎年 NII が取りまとめて CERN に支払い、各参加機関は NII に支払っています。

詳しい内容は Web サイトをご覧ください。(http://www.nii.ac.jp/sparc/scoap3/)

■ SPARC Japan セミナー報告

CC BY

第 2 回 SPARC Japan セミナー2017 (オープンアクセス・サミット 2017)

「プレプリントとオープンアクセス」

2017 年 10 月 30 日（月） 国立情報学研究所 12F 会議室 参加者：63 名

本セミナーでは、オープンアクセスの展開において不可欠な要素であるプレプリントサーバについて取り上げました。プレプリントサーバの機能と変遷を確認しつつ、研究者と出版者からの発表を交えて、研究推進への寄与、持続性のあるビジネスモデルの構築、質の確保などの将来の課題を含めて議論しました。

次ページ以降に、当日参加者のコメント（抜粋）、企画後記およびドキュメント全文（再掲）を掲載しています。その他の情報は SPARC Japan の Web サイトをご覧ください。(http://www.nii.ac.jp/sparc/event/2017/20171030.html)

概要

CC BY

オープンアクセスは、主にリポジトリを活用するグリーン OA と、ジャーナルを無料化するゴールド OA と呼ばれる方法のいずれかによって実現されると考えられている。arXiv を嚆矢とする分野別プレプリントサーバはこれまでリポジトリの一種別として位置付けられ、かつ、査読の前の草稿であることが原則であるため、研究成果の商業的流通への影響が少ない存在として理解されてきた。しかし、なかなかプレプリント流通が普及しなかった化学、生命科学の分野でも近年急速にプレプリント掲載数が増加する傾向が顕著となり、有力な学会誌をもつ国際学会(ChemRxiv の ACS 他)や商業出版者(BioRN の Elsevier)がプレプリントサービスの提供に参入してきている。

さらに、これらのプラットフォームはデータ公開を含めて設計されており、オープンアクセスの展開におけるプレプリントサーバの位置づけは大きく変化しつつあるといってよい。

今年度の第 2 回目の SPARC Japan セミナーでは、プレプリントサーバの機能、運営の約 30 年の変遷を振り返りつつ、この変化の渦中にある研究者からの提言、さらに研究者ネットワークの商業出版者にとっての意義を論じ、研究推進への寄与、持続性のあるビジネスモデルの構築、質の確保などの将来の課題を含めてこれからのオープンアクセスを展望する。

参加者から

（大学/図書館関係）
・プレプリントの価値が認められている分野がある、そしてこれから最も存在価値が上がりそうということをリアルに感じられました。
・arXiv と SSRN の話の対比は非常に興味深い。持続可能性についてさらに議論を聞きたい。

（企業/図書館関係）
・図書館側からすると、プレプリントサーバの存在は、まだ注目できていない気がします。今後も、アンテナを張っていかないと、研究者の方からのレファレンスに応えられなさそうです。

（その他/図書館関係）
・最近の SPARC セミナーは研究者さんの登壇が多く、実際に何が必要とされているのか感じ取ることができてとても良いと思います。図書館関係者だけだと研究の現場で求められていることから外れた独りよがりになりそうで、視野を広く持つことの大切さが身にしみます。でも、研究者さんの考えていること、必要としていること、困っていることを聞くたびに図書館が考えていることと距離を感じ、どうしたものだろうかと途方にくれます。研究者は機関を超えて分野ごとのグローバルなコミュニティでものごとを考えています。研究機関の図書館の役割はあるのでしょうか？必要とされる図書館になるには、どうしたらよいのでしょう？〔動画中継視聴者〕

（その他/研究者）
・状況確認、update に役に立った。研究者・論文執筆者視点（プレプリントを知らないシニアな研究者・学協会長などが研究者に／学協会に／各自所属機関にとって役に立つ可能性を感じるのかどうかなど）からの議論を深めてもよいのではとも思いました。

企画後記

CC BY

企画のプロセスで不馴れだったため周りに迷惑をかけてしまいました。しかしながら、結果としてプレプリントに関してさまざまな分野の研究者のみならず、SSRN の Gordon さんにも運良くお話ししていただくことが叶い、大変刺激的なセミナーになって良かったです。

坊農　秀雅
（情報・システム研究機構
ライフサイエンス統合データベースセンター）

今回 twitter 発信を担当しました。今回のテーマ“プレプリント”は日頃の業務ではあまり意識しておらず、大変勉強になりました。研究の速度が増していることがプレプリントを促進している、との言葉が印象的でした。各分野の研究者の具体的なお話を聞くことができ、プレプリントの分野ごとの浸透状況や出版社と研究者の関係性など、常に最新の動向を認識することの重要性を痛感しました。

笹渕　洋子
（早稲田大学図書館）

今回のセミナーは、プレプリントに重点を置いた企画でした。図書館職員の方にはまだ自分たちには馴染みが薄いと思われたのかもしれませんが、コーネル大学の図書館が arXiv.org を運営しているように、そう遠くない時期に図書館職員の仕事の一つとなる日が来るのかもしれません。オープンサイエンス時代の大学図書館の新たな機能のひとつとして、これからもプレプリントには注目していきたいと思います。

梶原　茂寿
（室蘭工業大学附属図書館）

第 2 回 SPARC Japan セミナー2017

「プレプリントとオープンアクセス」

開会挨拶/趣旨説明

坊農 秀雅

（情報・システム研究機構 ライフサイエンス統合データベースセンター）

CC BY

坊農 秀雅

理化学研究所においてFANTOM(Functional annotation of mouse)プロジェクトの立ち上げに関わった後、埼玉医科大学ゲノム医学研究センターを経て、2007年7月より情報・システム研究機構 ライフサイエンス統合データベースセンター(DBCLS)にて統合データベースプロジェクト(統合DB)に従事。統合DBの広報・普及活動として統合TVや統合データベース講習会AJACSの立ち上げに関わり、現在は国立遺伝学研究所にて日本DNAデータバンク(DDBJ)と連携して大規模塩基配列データ利用環境の構築とそれを利用した生命科学研究を行っている。京都大学博士(理学)。

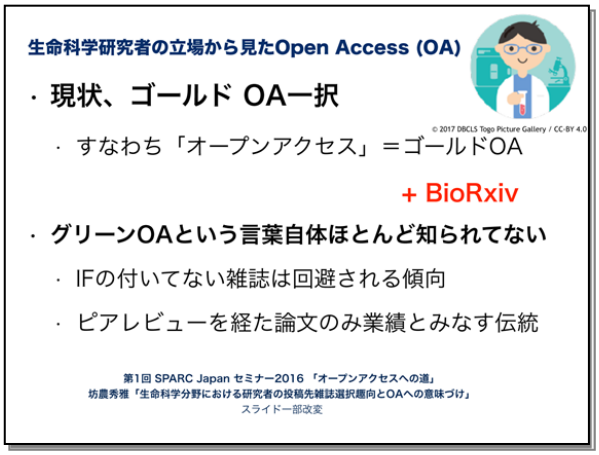
第 2 回 SPARC Japan セミナー2017「プレプリントとオープンアクセス」にご参加いただき、誠にありがとうございます。本日のセミナーでは、四つの講演の後、全体議論が行われます。プレプリントに関するセミナーはこれまであまりなかったので、議論して新しい知識を得られればと思います。

注目される bioRxiv

私は 1 年前の 9 月に行われた第 1 回 SPARC Japan セミナー2016「オープンアクセスへの道」で、生命科学において、研究者はどのような雑誌に投稿するかという話をしました。図 1 はそのときに出したスライドの 1 枚です。私自身は生命科学研究者で、生命科学研究者から見たオープンアクセスはどのようなものかと考えたときに、基本的には「ゴールドオープンアクセス」と呼ばれる、オープンアクセスの費用を論文を出した投稿者が持つという選択肢ほぼ一択で、オープンアクセス＝ゴールドオープンアクセスという状況でした。

「グリーンオープンアクセス」と呼ばれる、機関アーカイブなどに出すタイプのオープンアクセスはその言葉自体あまり知られていませんでした。その原因は、インパクトファクターと呼ばれる数値が付いてない雑誌には出したくないという研究者が生命科学系には多いことです。

もう一つは、ピアレビューを経た論文のみを業績とするという伝統があるからだということを昨年 9 月にお話ししたのですが、それ以降、急速に情勢が変わって、生命科学の分野においても bioRxiv というものが昨年の後半ぐらいからいわれるようになり、前から bioRxiv 自体はあったのですが、無視できなくなって



（図 1）

きました。余談ですが、これは「バイオアーカイブ」と読みます。まだあまり広まっていないのですが、かなり注目されてきています。bioRxiv に挙げられた論文の中で、この 1 年で大きなインパクトがあったと思うのは、図 2 です。

最新の次世代 DNA シーケンサー「HiSeq 4000」に不具合がある、データがおかしいのではないかという論文が今年 4 月にピアレビューを経ずに bioRxiv にアップされ、われわれの業界では非常に問題になりました。実際にこのシーケンサーから出てきたデータの幾つかはおかしいのではないかと騒がれていて、現状でもこれはまだ解決していないのではないかと思うのですが、そういう状況が非常に素早く研究者に伝わったということで、このプレプリントは非常に注目されるようになってきて、それ以外でも最新の機械に対する評価は bioRxiv で発信されるということが起きています。

今回のセミナーの準備をしている間にも、「Science」誌に、プレプリントに関する記事が幾つか出ていました（図 3）。エディトリアルにおけるプレプリントのエコシステム、プレプリントとどう付き合っていくか、「Science」はそれをどう考えるかというようなことが書かれているので、これも図書館に行って読んでみてはいかがかと思います。ここでは詳しくは述べません。

私自身の専門である生命科学分野の話ばかりしましたが、今日はそれに限らず、プレプリント全般に関してお話をいろいろな分野の方からしていただきます。

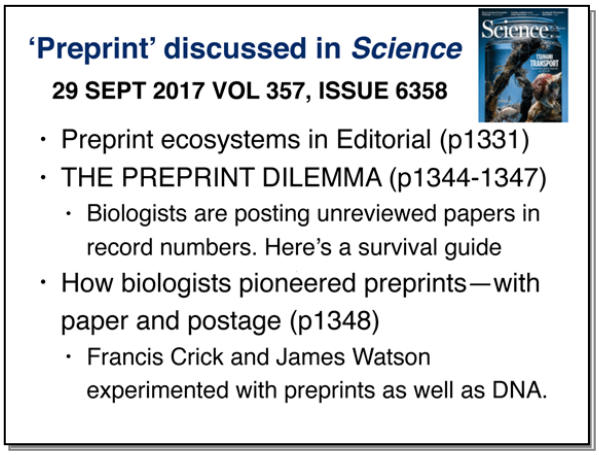
（図 2）

プレプリント自体や arXiv というプレプリントサーバはそれこそ約 30 年前から進められています。生命科学分野で急に注目されているので、そこに出そうという人も出てきているということです。

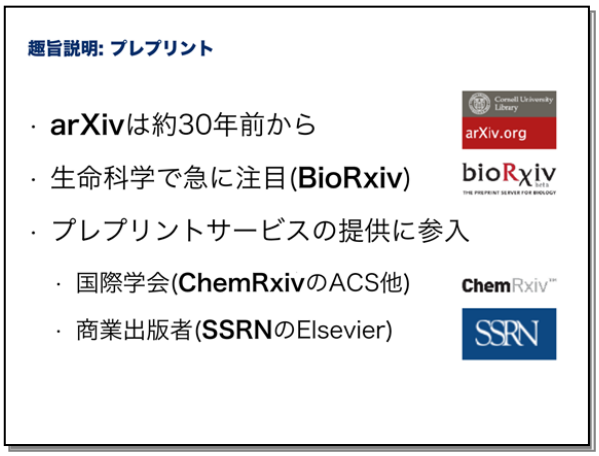
それ以外の分野においても、例えば化学の分野においても ChemRxiv というプレプリントのサービスが公開されています（図 4）。商業出版社においても、Elsevier が SSRN というリポジトリで参入しています。

本セミナーの趣旨

本日は、まずプレプリントサーバに関わってこられた引原先生にお話しいただきます。そして、変化の渦中にある、化学分野と生命科学分野の研究者お二人に、普段されている研究活動も含めてお話しいただき、その中でオープンアクセスの今後についても話し合いたいと思います。SSRN のマネージングディレクターで



（図 3）



（図 4）

ある Gregg Gordon さんにもお話を伺って、議論を進めます。

この会場の内容は YouTube live で配信されていますし、質問事項等は Twitter で発信していただけます。ハッシュタグ（#sparcjp201702）を付けていただくとわれわれも見つけやすく、こちらで質問にお答えすることもできると思いますので、ネットで見ている方は Twitter でハッシュタグを付けて、どの先生への質問かということも明記していただいて発信していただければと思います。

第 2 回 SPARC Japan セミナー2017

「プレプリントとオープンアクセス」

arXiv.org の次世代システムの公開と戦略

引原 隆士

（京都大学図書館機構長/arXiv.org 会員コンソーシアム代表）

講演要旨

arXiv.org は、1991 年にロスアラモス国立研究所で物理学分野のプレプリントサーバとして設立され、その後、数学、非線形科学、計算機科学、統計学、計量ファイナンス等へ領域を広げた。2011 年に Cornell 大学図書館が運営している。本年、新たに電気システム科学、計量経済学のカテゴリーを増やし、次世代システムの概要を固めた。ボードメンバー会議から、新システムのアイデアと arXiv.org の今後の戦略について情報提供する。

引原 隆士

1987 京都大学大学院工学研究科電気工学専攻博士課程修了。2001年京都大学大学院工学研究科教授。京都大学工学博士。この間コーネル大学客員研究員。2012より 京都大学図書館機構・機構長・附属図書館長。現在、内閣府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委員、文部科学省第9期学術情報委員会委員。

私はもともと arXiv.org（アーカイブドットオルグ）のユーザーであり、また、Member Advisory Board（MAB）という arXiv.org のアドバイザリーボードに、国立情報学研究所から選出されて、昨年度からメンバーとして参加しています。

arXiv.org がカバーしている分野は主に数理物理ですが、最近はバイオロジーが加わったり、さらに新しい分野が追加されています。こういう分野について arXiv.org が何をしてきたかということをお話しすると同時に、今後どういうことを目指しているか、私の知る限りでお話ししたいと思います。YouTube に、arXiv.org がどういう歴史をたどってきたか、何をしてきたかが出ているので、参照していただければと思います（https://www.youtube.com/watch?v=ntoxZzh0ha8）。

arXiv.org のこれまで

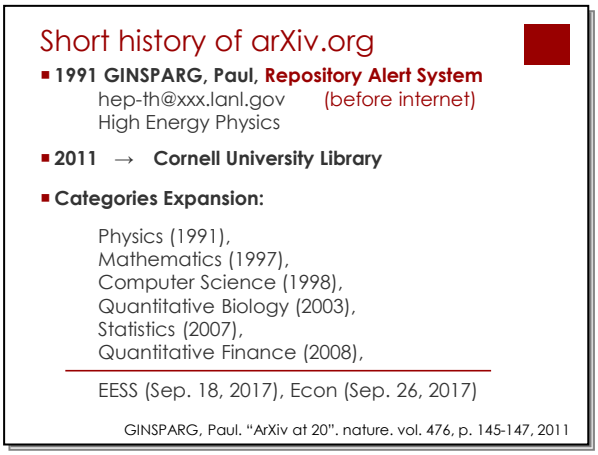
arXiv.org の創始者は Paul Ginsparg で、ロスアラモス国立研究所の彼のデスク上の HP のサーバから arXiv.org は始まっています。それが 1991 年で、インターネットが始まる前です。私が在外研究に行ったのが 1993 年で、その直前は BITNET などが連絡手段でした。その時代にメールベースで始められたのがそもそもの始まりです。

それをまとめると図 1 に書いてあるとおりです。高エネルギー物理学の論文のやりとりを創始者の Ginsparg が始めました。お互いに論文をやりとりして、共有して、必要なものをこのメールベースのデータベースに早く入れるようにする、あるいはそれをアーカイブして置いておくということが行われました。

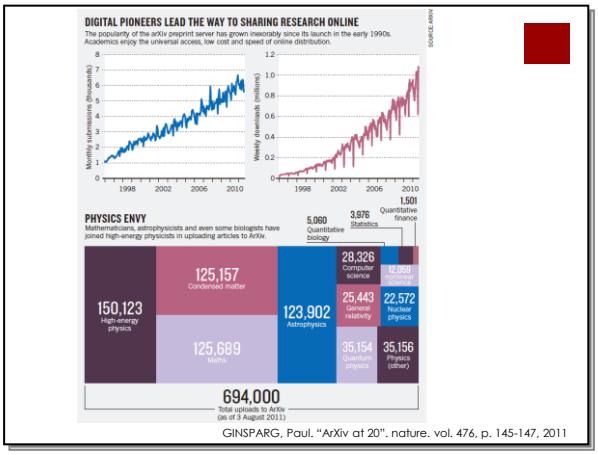
それから 20 年たったことをきっかけに、2011 年にコーネル大学の図書館に移行しました。私自身が在外

研究に行っていたのがコーネル大学で、それは 1993～1994 年ぐらいでした。その当時はライブラリはあまりこういうことを意識していませんでした。昨年、コーネルライブラリのミーティングに久しぶりに行って、どういう形でここに移行したかという経緯などもいろいろお聞きしました。それが今、エコシステムというか、そもそも始められたシステムをどうこれから発展させていくかという大きな契機になっていると思います。

分野については、1991 年に高エネルギー物理学の分野で始まっています。その後、1997 年に数学、1998 年にコンピュータサイエンス、2003 年に quantitative biology、2007 年に statistics、2008 年に quantitative finance で始まり、ここでいったん発展は止まっています。ところが、今年 2017 年 9 月 18 日に EESS（electrical engineering system science）が加わりました。



（図 1）



（図 2）

さらに、economics が 1 週間後の 26 日に加わっています。

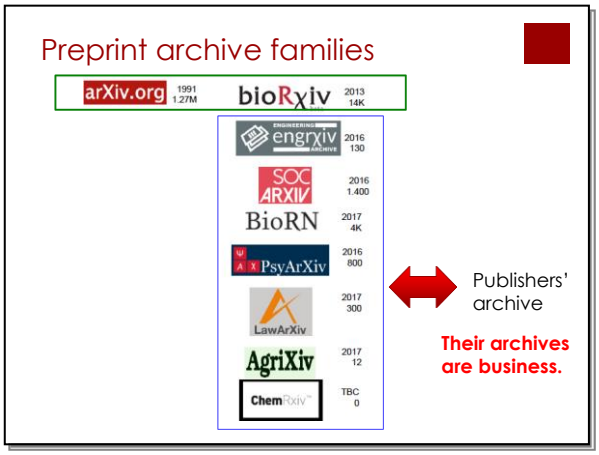
物理で始まって、物理の関連で次に数学が始まったというのは分かりやすいのですが、この後どのように発展していくかということを決めて、現在に至っているかというのが重要なキーになります。

図 2 は 2011 年に 20 周年を記念してまとめられたものです。20 周年の時点では、月ごとのサブミッションの増加は左上のグラフのようになっています。週ごとのダウンロードの増加は右上のグラフのようになっています。指数関数的に増えているということが見て取れると思います。

このダウンロードの対象あるいはサブミッションの対象になっているのが下のマップに示されている分野です。このような比率で入っていたのが 2011 年です。

今後これがどう変わっていくかというのが arXiv.org の中では非常に重要な観点です。これから広がっていく分野を加えつつ、それまでの分野を arXiv.org として維持していく、その流れをどうやって判断していくかというのが大きな課題になっています。

プレプリントサーバとしては、arXiv.org は 1991 年に始まりましたが、bioRxiv は 2013 年に始まりました。それ以外にも図 3 のようなものがあります。arXiv.org と bioRxiv は今、かなり連携しています。お互いの運用に関しても情報交換しており、これが今の bioRxiv の発展をサポートしていると思っていただければいいと思います。



（図 3）

ソーシャルサイエンスの SSRN など、パブリッシャー側に移ったアーカイブもあります。パブリッシャー側に移ったアーカイブは、もうビジネスに移ったものと見ています。ですから、研究者側にあるアーカイブ、ビジネス側にあるアーカイブの二通りあるということです。これを同一に話ができるかというと非常に難しい面があります。

arXiv.org は昨年 2016 年からドネーションを始めています。ドネーションのページがあり、ダウンロードする側とサブミッションする側、両方のユーザーに、このシステムを維持するためのドネーションをお願いしております。これがかなり大きな額で、arXiv.org を支えています。

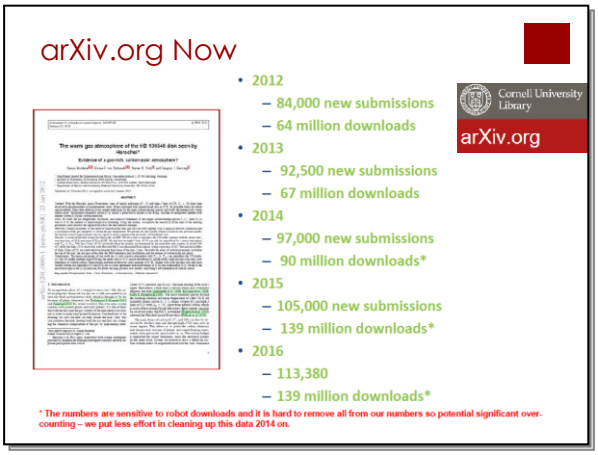
統計で見る arXiv.org

arXiv.org のサブミッション数は 2012 年に約 84,000 で、それからどんどん増えていって、2013 年に約 92,500、2016 年に約 113,380 です。ダウンロード数は 2016 年に 139 百万です（図 4）。ですから、膨大な数のダウンロードが始まっているのです。

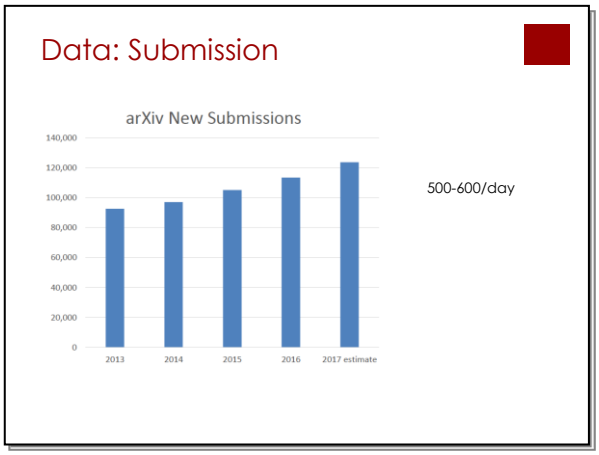
中にどれだけ面白いものがあるかということが当然重要ですが、arXiv.org の趣旨は、パブリッシュされる前の論文ができるだけ早い段階でユーザーに伝えられるということです。今はこれをいろいろな検索エンジンがウォッチしていて、いろいろなところからこの情報が上がったということを出せるような仕組みに変わりつつあります。ですから、当初のメールベースのもの、それからインターネットが始まったときにこのサイトにアクセスしたら見られたというものから、検索エンジンを介して提供するという状態にだんだん移っています。

図 5 はサブミッションの変化です。2017 年現在で約 12 万あります。1 日に 500～600 サブミッションされます。これを処理しなければいけないのです。この処理をコーネル大学のグループが今行っています。

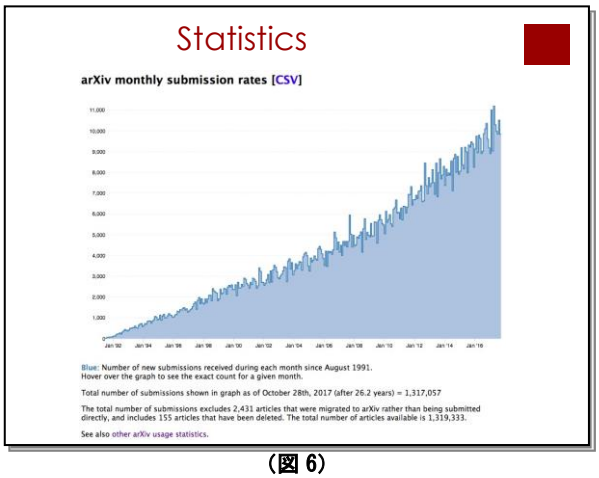
それをもう少し統計的に見ると図 6 のようになっています。1992 年は微々たるものだったのですが、現在はここまであります。サブミッションは、日本とは違い、夏休みはやはり少なく、クリスマスも少ないです。この波が大体毎年同じような傾向をたどっています。これは増加を総量で書いているので、その年々で出すと大体フラットな変化をしています。加算してリニアに増えています。



（図 4）



（図 5）



（図 6）

これではよく分かりませんので、もう少し分野ごとに見てみます（図 7）。左はトータルの数で見たものですが、割合に変えると右のようになります。当初は高エネルギー物理学だったので青色のところがほとんどだったのですが、それが時がたつにつれて大きく変わっていきます。Condensed matter physics が増えている時期もあります。Astrophysics、condensed matter physics が 2000 年あたりに高エネルギー物理学と同じぐらいの比率になります。その時期に数学が少しずつ増えていき、現在では高エネルギー物理よりも数学の割合が高いです。総数が増えているから、総数の中の数で見ていただかなければいけないのですが、他の物理やコンピュータサイエンスも増えていることが分かります。

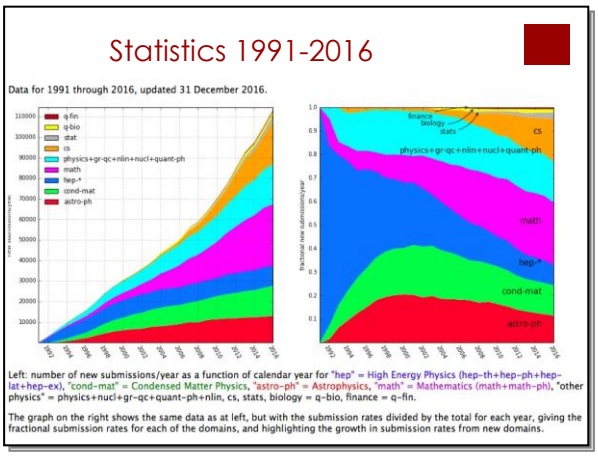
この変化を見ながらサーバとしての運営をしていて、例えば mathematics の中でも細かいカテゴリーを付けていく、あるいは condensed matter physics の中でも細かいカテゴリーを付けていくことによって見えやすいものに変えていくというような変化をさせています。

このうち、サブミッションとダウンロードを行っている国の内訳は図 8 です。まず、arXiv.org を支えているアメリカのコンソーシアムが 18%、アメリカのノンコンソーシアムが 21%と結構多いです。一つ一つの大学や研究所がコンソーシアムとして組んでいる部分と、そうでない部分が大体同じ比率になっています。UK が 7%、日本が 7%と非常に多いです。この中でも研究大学といわれるところが多くの論文をサブミッション・ダウンロードしています。スウェーデン、オランダも結構数が多いです。ドイツが面白くて、物理系で投稿するときには自動的に arXiv.org に上がるようなシステムを組みつつあるので、だんだん割合が増えているというのが現状です。

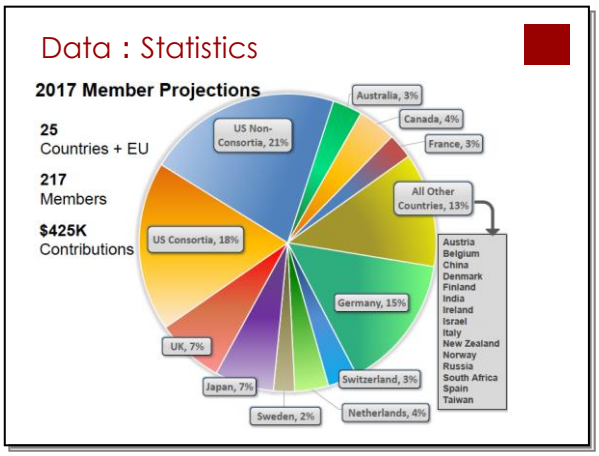
アジアでも増えており、中国のユーザーの割合が非常に高くなっています。また、投稿よりもダウンロードの数が非常に多いということが知られています。このような割合になっています。

これを幾つかのサーバで支えています。以前は日本にもサーバがあり、京都大学の基礎物理学研究所にミラーサーバを置いていたのですが、それは約 2 年前（2015 年 12 月）廃止になって、現在は 14 のサーバがあり、そのうち五つのミラーサイトが中国、ドイツ、インド、スペイン、ロスアラモス国立研究所にあります。これらの国からアクセスするのが非常に難しい場合があるために置いています。日本の場合、ネットの回線が非常に強くなっているので、あまり問題なくなったのですが、中国やインドなどはミラーサーバを置くことによってアクセスがスムーズにいくようになりました。ドイツに関しては東ヨーロッパ系のところをサポートしているような形でしょうか。

もとのサーバシステムもまだコーネル大学の中にあります。以上のシステムで今は運用していますが、これをそれぞれの国でスタンドアローンでマシンの上に置くだけではなく、今は Amazon のウェブサーバのクラウドに移しつつあります。Amazon に移してしまう



（図 7）



（図 8）

ことがいいかどうかというのは議論がありますが、移して、よりアクセスがしやすいものにする方向で動いています。

問題は、サーバはいいのですが、これのデータベースの部分とどうやってリンクを維持していくかということです。先ほど、データが非常に増えているということをお見せしましたが、そのデータベースに全部リンクしないといけないのです。それをやるのは、マシン間では非常に難しいということが明らかです。

ですから、今、クラウドの上に持っていこうとしているのですが、そのときにやはり Amazon だけでいいのかという議論はされています。現状は様子を見るということで、Amazon のクラウドに移すという方向で進めていますが、今後もこれがずっと続くかどうかというのはまだ分かりません。

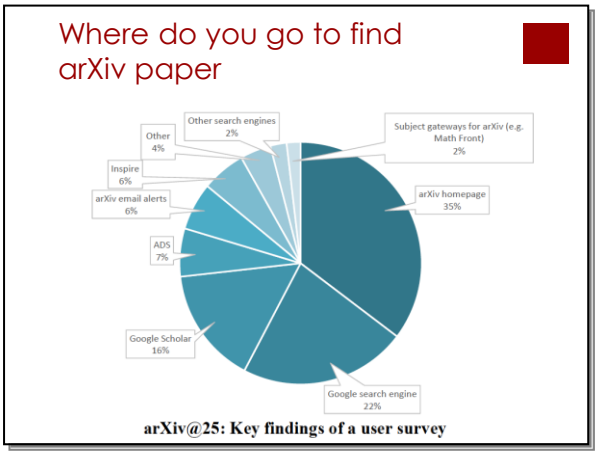
次は、どこから論文にアクセスしているかです（図 9)。最初はサーバに直接アクセスして、そのサーチを使って論文を見るという形だったのですが、2 年ほど前のデータでは、本体の arXiv.org のホームページからのサーチ以外に、Google のサーチエンジンを使っているもの、Google Scholar を使っているもの、ADS を使っているものがあります。それから arXiv email alerts とありますが、そもそも、arXiv.org は論文をアップしたときに、アップされたというカテゴリーごとの e メールをアラートで送ってくるシステムだったので、そのシステムがまだ生きています。今、他のところでも同じようなシステムを使っています。「あなたの論文が引用されました」というようなアラームがあると思うのですが、それが結構有効だということがよく分かります。

INSPIRE は別のデータベースサーチになりますが、こちらのリンクも結構強くて、CERN などからこちらを経由しています。これは重要なデータで、どこにどのように新しいデータベース、新しい arXiv.org が、システムを組んでリンクを付けていくとサーチが進むかというのが分かりやすいデータになるかと思います。これが状況です。

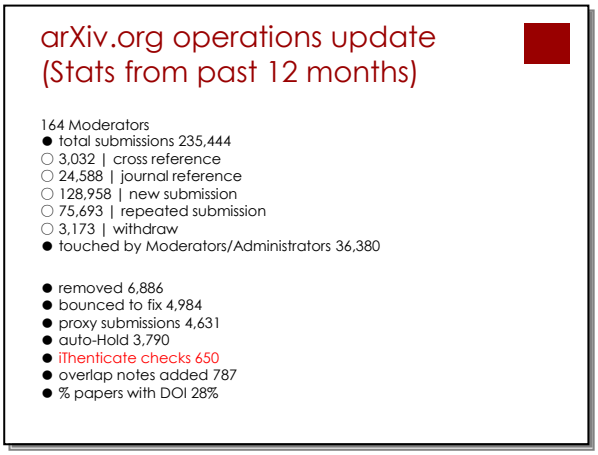
少し中身に入ります。先ほどトータルサブミッションの話をしました。そのトータルのサブミッションは約 230,000 あるのですが、そのうち、ジャーナルのリファレンスが約 25,000、クロスリファレンスが約 3,000、ニューサブミッションが約 130,000、再投稿が約 75,000、withdraw（取りやめ）が幾つかあります（図 10)。

これらをアドミニストレーターが処理するのですが、非常に大きな数を扱っていることが分かるかと思います。こういうことをどのような人たちがやっているかは、オープンになっている情報からご覧いただければと思います。

この中でリムーブされたものが約 6,800、それから、特に気を付けて見ていただければいいかと思いますが、iThenticate、要するにチェッカーで約 650 がリジェクトされています。だから arXiv.org のユーザーも決して紳士ではないということがよく分かります。



（図 9）



（図 10）

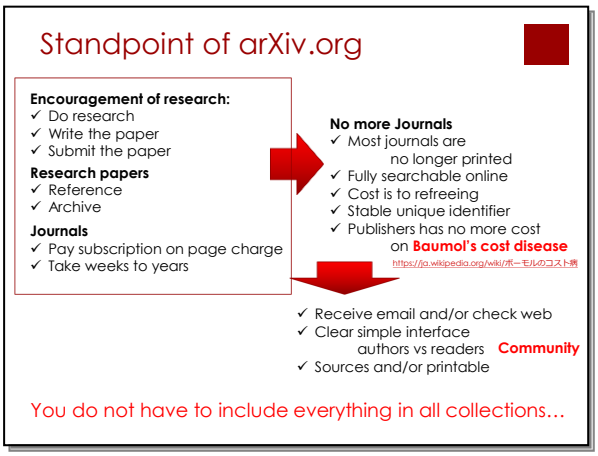
iThenticate で引っ掛かるというのはよほどのものだと思いますが、こういうものが出ているということです。これには説明のとき、みんな苦笑いしていました。

それから、DOI を付したものが約 28%あります。このプレプリントでも DOI や ORCID などを著者のメリットのために使うという考え方をしています。

arXiv.org の目指す方向

研究をエンカレッジするというのが arXiv.org の基本的な考え方です（図 11）。研究してください、論文を書いてください、投稿してくださいとエンカレッジしているわけです。そのときにリファレンスのペーパーを見えやすくして、アーカイブもきちんとしてあげましょうというのが arXiv.org の考え方なのです。

一方、arXiv.org が何をメリットとして主張するかというと、当たり前のことですが、ページチャージは要りません。それから、論文を投稿して掲載されるまでに何年もかかからず、1～2 週間でプレプリントが上がるということです。化学系や物性系だとすぐに掲載されることが結構あり、何年もというのが皆さんにあまり伝わらないケースが多いのですが、システム系や数学系、物理でも理論系だと、1 年たっても 2 年たってもまだ査読中ということが結構あります。私も若いころに 1～2 年かかるということがあり、このままいくと私は首になるというぐらい、たくさん出しているのですが、査読が終わったものが何も出てこないのです。どんな論文を出しているのかも、著作権の関係で出せないのです。



（図 11）

そういうものに対して、出したということをきちんと arXiv.org でオープンにして、それが伝わるようにして、物理なら物理の世界が早く進むようにエンカレッジしていくというのが基本的な考え方です。ですから、やはり物理特有の状況に根差してこうなってきたというのがよく分かると思います。

こういう立場で始まったのですが、現在、運営者たちはこう言います。「もう論文は要らない。どんな論文ももうプリントされていないではないか、ほとんどがオンラインでサーチできるようになっている、コストはレフェリーのお金にすぎない」と。レフェリーは研究者がしています。よく日本でも産地直送と仲買を経た場合の値段の違いという話がありますが、産地直送でいいのではないかという話になるわけです。

パブリッシャーは、「ボーモルのコスト病」といって、もはやもうける仕組みにはなっていません。「ボーモルのコスト病」というのは Wikipedia を引いていただいたら分かると思うのですが、増やすことによってもうかるシステムではない、本来は価値で動いているのにそれを数で処理している、運営者たちは、そういうシステムの中に入る必要はないということを言っているのです。

ではどうしていくかというと、e メールを使うか、ウェブを使うかは別として、コミュニティをもっとクリアにしていく、コミュニティ間のやりとりをエンカレッジしていくような方向にして、著者とリーダーが直接やりとりできるシステムとして維持したいということです。それから、ソースはプリンタブルにしておくということができれば、自分たちが好きなように使えるということになります。

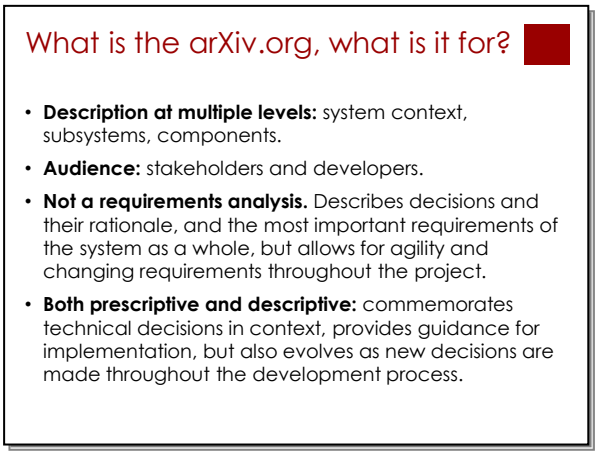
多くの場合、システムが大きくなってくると、もっと入れてください、この分野も増やしてくださいなどいろいろ言われると思うのですが、全ての分野を入れるつもりはなく、自分たちの考え方で維持できる分野だけを増やしていこうとしています。そういう共通認識を持っている分野を探し出して動かしているという

ことです。

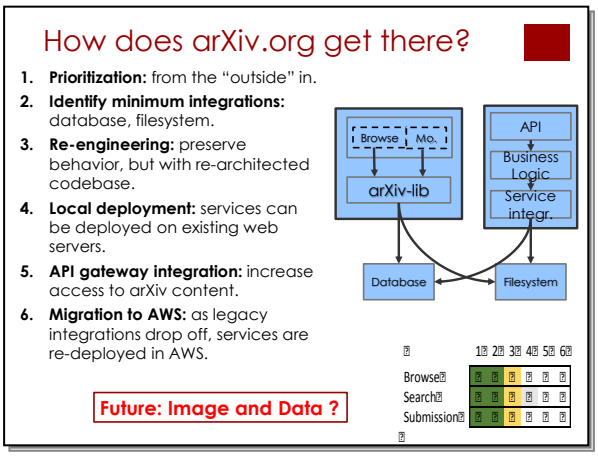
arXiv.org は今、次世代に向かって動いています。図 12 にはステークホルダーとデベロッパーと書いてあります。そのステークホルダーはリーダーであり、サブミッションするオーサーであったりするわけですが、それ以外にパブリッシャーも当然います。デベロッパーというのは、このシステムをつくっている人たちも含まれているのですが、その中で、どういうシステムを次の世代でつくっていったらいいかという議論を非常に熱心に行っています。これが恐らく次の世代をつくります。

従来のシステムは何でもかんでも入れ込んでいるデータベースでした。これはもう今の時代では成立しません。ですので、それを書き換えて、現在は application programming interface（API）を使って、データベースと API をフィルタリングするような形を取って、最終的には図 13 のような形を目指します。ブラウザとライブラリがあって、API が管理する business logic と service integration のプログラムを置いておいて、それでデータベースとファイルシステムをやりとりするというようなことです。これらに対して従来型のプログラムをそのまま延長するのではなく、Python や Flask web プログラム、Docker や REST APIs など、最近使われているものをフルに取り入れてセキュリティを保ちながらやっていくということを計画して進めています（図 14）。API に持っていくことによって何が良くなるかというと、今までのデータベースだけではなく、他のものも取り込める可能性が出てきます。そのために、将来必要なイメージや raw data といったものへのデータベースのリンクも可能になるであろうことが予想されます。

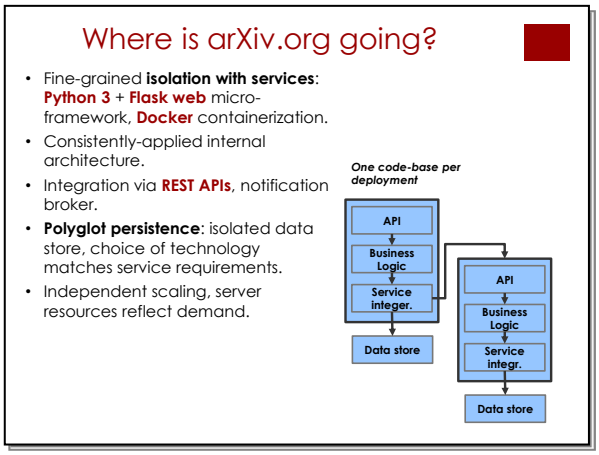
実際の絵は図 15 のようになります。arXiv.org のシ



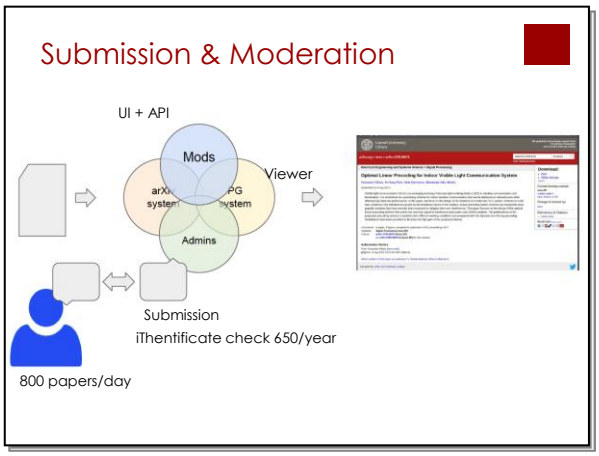
（図 12）



（図 13）



（図 14）

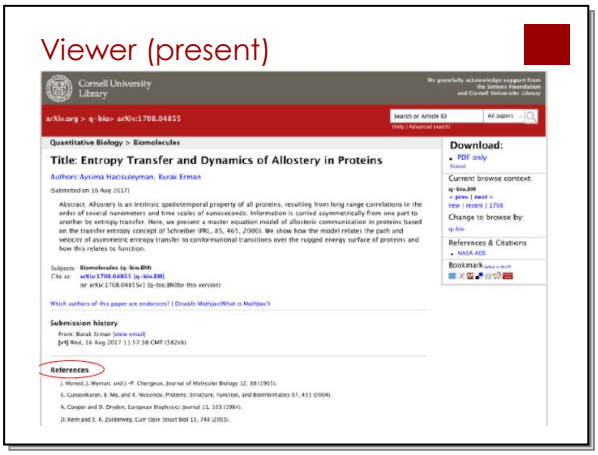


（図 15）

ステムがあって、サブミッションがあったときに、ビューワー側とアドミニストレーターが見るサブミッション側の処理と、モデレーターがお互いを見ながら全体のシステムを運営していくという流れをつくっています。

ビューワーについては、今までは図 16 のようなビューワーで使いにくさがあるということを皆さん分かっていて、このビューワーに対して他の方法を示しています。

例えばリファレンスだけを別で見られるようなものです（図 17）。こういうのはよくあります。このように変えたのですが、アンケートを取った結果、前の方がいいという人が結構多いのです。アンケートの結果、49%が前の方がいいと答え、51%が新しい方がいいと答えました。要するに、ビューワーは慣れたものがいいということがよく分かります。



（図 16）

（図 17）

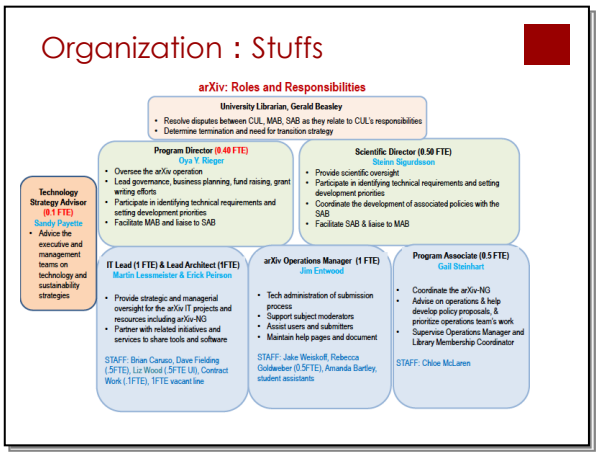
システムの運用は図 18 のようになっています。リーダーグループがあって、私が属している Member Advisory Board（MAB）というアドバイザリーボードがあって、Scientific Advisory Board（SAB）が開発をしています。それから、コーネル大学のライブラリのアドミニストレーションがサポートしています。

これを非常に少ない人員で行っています。図 19 の資料もオープンなものです。full-time equivalent（FTE）ですから、1 日を 1 とすれば、0.1〜1 のエフォートで、エフォート管理をしながら全体を運営しています。ライブラリとこちらの仕事を共有しながらやっているというのが現状です。

パブリッシャーとの関係

ここが今日一番重要かと思いますが、パブリッシャーというものをどう取り込むか、どう連携するかとい

（図 18）



（図 19）

うのは昨年かなり議論がありました（図 20）。パブリッシャーはしばらく置いておこうという結論です。それに対して、学会に関してはもう少し協調していこうと結論が昨年度出て、今年度は実際にディスカッションを行っています。物理系ですから APS や AIP、ACS、AMS などのエディターあるいはトップの方々に来ていただいて議論しました。もうすぐ IEEE も運営の中で一緒にやります。

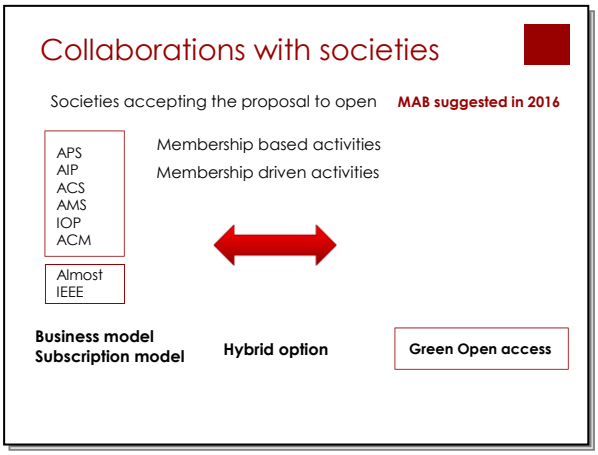
プレプリントですから、出しているということをオープンにするのは構わない、オープンになった後、実際にパブリッシュされたものと共存するということも認めるというように、連携しながらお互いにハイブリッドの形で生きていこうとしています。

その結果、狙うところは、グリーンオープンアクセスです。パブリッシャーはゴールドが多いですが、学会は主にハイブリッドです。ですから、ハイブリッドを維持することで、arXiv.org 経由で動いていくというのも非常にアクセスが増えるという意味でいいと認識されているようです。

今、申し上げたものについてはお金が当然必要なので、これはスローン財団がサポートしたという記事が arXiv.org に出ていました（図 21）。

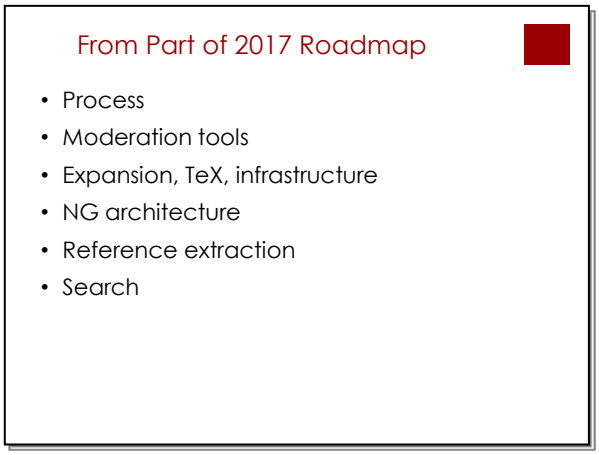
2017 年以降のロードマップはまとめて書いておいたので、参考にご覧ください（図 22）。

今いろいろな取り組みをしていますが、ユーザーインターフェースと API で、全体のデータベース、アクセスのビューワーなどを再構築していて、それによって他との連携ができるような形につくり替えようとしているとご理解いただければよろしいかと思います（図 23）。また、シングルクリックで全てのオペレーションができるような形にするという再構築もしています（図 24）。

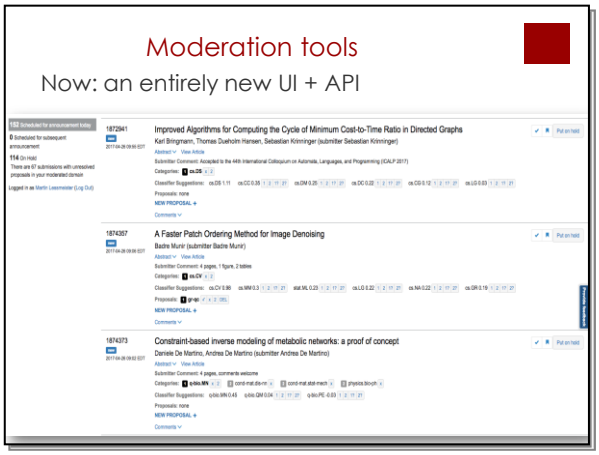


（図 20）

（図 21）



（図 22）



（図 23）

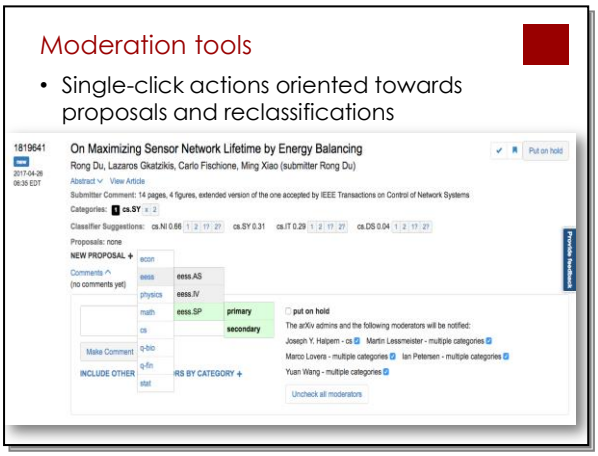
TeX システムもバージョンアップしています。arXiv.org は PDF で出すこともできますが、TeX のソースからアップすることも当然できます。物理系は TeX でやっているケースが多いので、それがそのまま使われています。

図 25 はシステムのハイライトです。いろいろなソフトを使っています。ネクストジェネレーションのアーキテクチャは、まとめると図 26 の 5 項目になります。インテグレーションが非常に重要なのですが、その前のテクノロジーを現代風に書き換えるということをやっているので、それによってセキュリティの問題が当然起きてきます。セキュリティの問題に関しては今もやりとりをしていて、チェックしているという状況です。ですから、完全に確保できたという状況ではないと思っています。

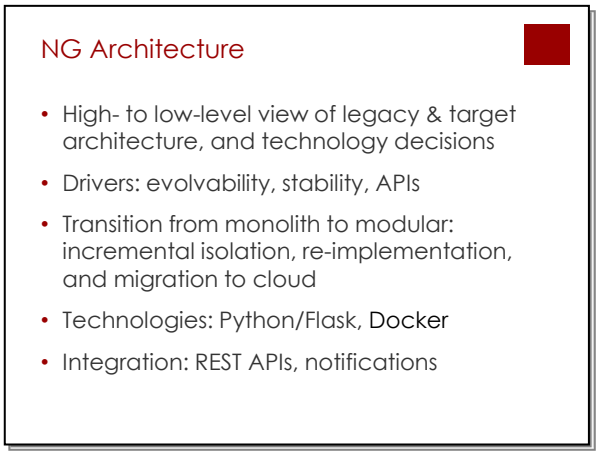
図 27 は、全体の運営としてどういうサステナビリティをどういう考え方でやればいいかというマップです。これは Oya Rieger さんが全体の運営の中でまとめたものです。バランス良く全体を運営したいと主張しています。そのソースになる、ファンドによるお金、各コンソーシアムが支払うお金、大学が支払うお金をどう当てはめていかないといけないかが描かれています。本当は予算の関係をお見せしようと思ったのですが、ビデオで流れるので今日は割愛させていただきました。ウェブで見られますので自分で探してみてください。

結論

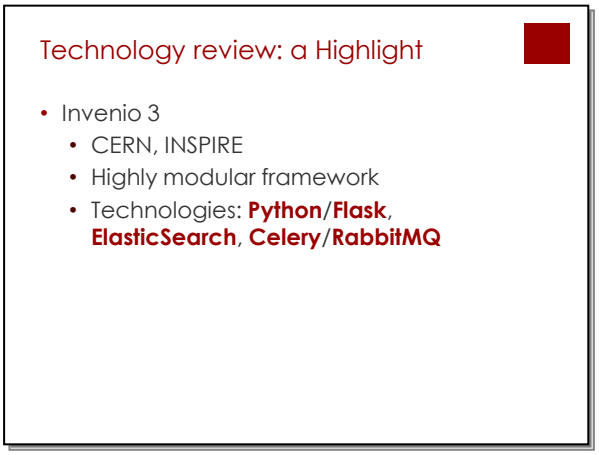
オープンライセンスに対するライセンスは何かという議論がずっと続いています。また、エンバーゴした後、ハイブリッドで出しているものと、もともとのパブリッシュされたものとの関係がまだ微妙なところが



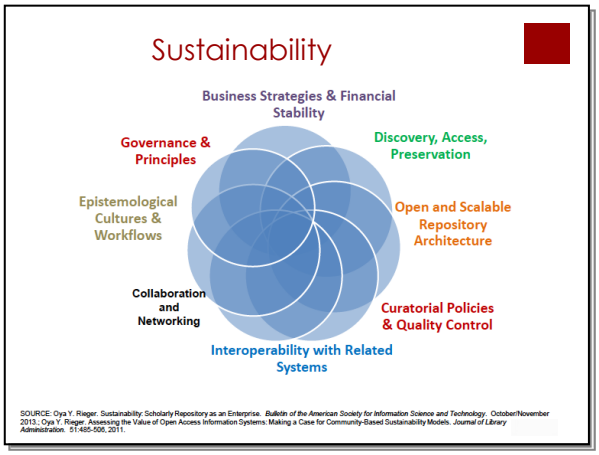
（図 24）



（図 26）



（図 25）



（図 27）

あります。エンバーゴしているからいいのではないかと思われますが、何が二次的にという話になったときに、どちらがソースかというのが分からない面があります。そういう部分が残っています。

arXiv.org の話の中でこれは共通認識だと思うのですが、別にパブリッシングモデルを変えたいと思っているわけではないのです。ユーザーを変えたいのです。ユーザーが変わってくれればシステムはもっと発展するし、今のパブリッシャーから多く要求されている支払いを抑えることもできるはずです。パブリッシャー相手に訴訟問題を起こしても仕方がないので、ユーザーが自発的に変わっていくようなシステム、それがやりやすいようなシステムをつくり上げていくというのが今からの流れです。ネクストジェネレーションのシステムもこれが変わるものではありません。ですから、これからパブリッシャーとして arXiv.org が出ていくというわけではなく、やはりコミュニティがどうやって自分たちを守るかという考え方に根差していると思っていただければよろしいかと思います。

機関リポジトリや他のアーカイブと協調していくことが必要だというのは認識されていますが、例えばリポジトリ関係は、日本の考え方とアメリカの考え方、ヨーロッパの考え方が全く異なります。日本のリポジトリはどちらかというとプレプリント的なものになっているので、既にアーカイブになっているのですが、アメリカの場合はそうではなくてビデオや講義録などが中心です。ですから、一括して同じようにはできないということが言えます。

逆に言えば、日本の場合、リポジトリは既にプレプリントの機能を持ちつつあるということです。ですから、それを発展させることによって、うまい連携ができる可能性はあると思います。そういう議論も今年提案させていただいたのですが、やはり国情によって考え方が違うので、プレプリントサーバとリポジトリの関係は日本なりのものを構築していく必要があるだろうというのが結論です。

●**フロア 1**　国立情報学研究所の武田です。今後、システムをサステナブルに維持するということに関して、資金はファウンデーションの資金、コーネル大学の支援、コミュニティ（事実上は各大学等）の支援から成っており、一方で 2017 年も分野を広げるということですが、コストと、分野を広げるということについての議論はあるのでしょうか。

●**引原**　非常に微妙な質問なのですが、財団からの資金は主には次世代システムの開発に回っていると思っていただければとよいかと思います。ですから、今年は昨年から予算が倍増しています。新しいシステムと旧来のシステムを並列して走らせなければいけないということがあって、そのようになっています。

今はシステムを維持するランニングコストは利用者が支払っています。日本でも NII をベースにコンソーシアムが支払いをしているのですが、今、払っているところが利用者全体の 80%程度にとどまっています。残りは、これから増えてくれば当然徴収すると思います。

それ以外のやり方は、先ほどミラーサイトについてお話ししましたが、そういう運営で頑張っていただくということです。日本も昔そうだったと思うのですが、そういう形を取って、何らかの貢献をしていただくような考え方になるかと思います。

最後に、分野を広げるという話については、分野はなかなか広がらなかったのです。今はほとんど 10 年ぶりに広がったわけですが、これは旧来のカテゴリーの中で違う分野が起きてきているからです。それを外出ししたと思っていただいた方がいいかと思います。EESS というのは「EE」と入っていますが、ほとんどはシステムサイエンスなので、バイオロジー関係でもシステムサイエンス的なものは結構ありますし、そういうものが今までのカテゴリーの中ではユーザーにリンクしにくい部分があったので外出ししたということです。外出ししたと同時に、もう少しそれに関連する分野を増やしていっています。エコノミクスも同じよ

うに、旧来あった分野を外出ししている、そういう形で少しずつ広げていますので、ドラスティックに変わったわけではないと思います。

●フロア 2　NICT の研究者です。私はこの分野に素人で、学会でプレプリントサーバの話を考えないかと言われて少し興味を持ちはじめたのですが、実は地球物理学でアメリカ地球物理学連合（AGU）がプレプリントサーバを今立ち上げているのです。arXiv.org と少し微妙な関係のようなのですが、日本の日本地球惑星科学連合に対して一緒にプレプリントサーバをやりませんかという声を掛けてきました。日本側は、地球物理はプレプリントサーバはなじみがなくて、私もそうなのです。

先取権やジャーナルで受理するかなど、いろいろな問題があるかと思うのですが、そういうのを学会でオーソライズして、われわれは先取権をプレプリントサーバ、ペーパーに認めますとか、われわれの発行するジャーナルは必ずプレプリントからでも投稿を認めますといったことを学会長自ら全員が承諾して、そのコミュニティがプレプリントサーバを受け入れないと、うまくいかないのではないかという雰囲気もあるのです。実際に今動いているプレプリントサーバの上で投稿されている方々の分野では、コミュニティの規範、norm、カルチャーが既に確立していて、業績をそこで出していくようなプロセスが動きはじめているのでしょうか。

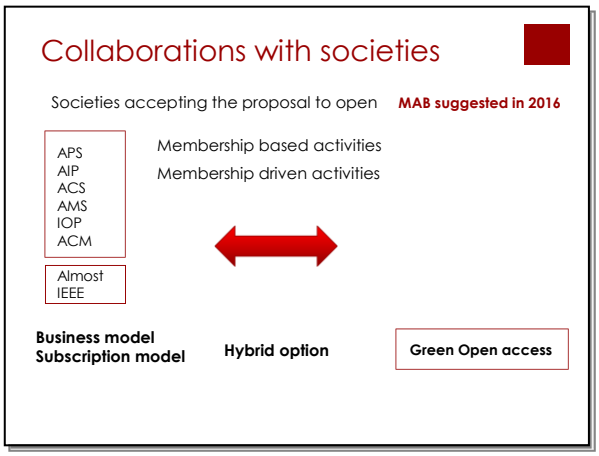
●引原　それもまた微妙なご質問だと思うのです。やはり分野によっても大きく違います。高エネルギー物理学では、SCOAP³ の話もありますが、共有化、早く出す、グリーンオープンアクセスというのが文化として定着しているのだと思います。それをコミュニティが支えているという現状があります。

一方、先取権ということを非常にセンシティブに捉えている分野と、2〜3 年もかかるのだったら早くコミュニティに帰する方がいいと考えている分野と、まちまちです。ですから、物理の分野の考え方がどこまで他に適用できるかというのは、色合いを見定めなければいけないだろうと思っています。

新興のプレプリントサーバの危ないところは、ある程度データがたまってくると出版社が目を付けます。そのときに規模の問題があって、小さいところは簡単に扱えます。M&A と一緒です。arXiv.org ぐらいになると、もう少し扱いきれない部分があります。クラウドになったら分かりませんが、そういう部分でやはり定着したというところの考え方を維持できれば、まだ大丈夫だろうと思っています。

コミュニティというのは、学会が残っている国は学会が持っています。そういうところは、出版を持っていても、基本的にプレプリントで出して先取権を任せて、あとは学会が最後のピアレビュー後の出版を受け持つのだというお互いの信頼関係を築くことが重要だろうということで、先ほどお見せした 6 学会（図 20）と議論してアプローチしているという状況だと思います。

●フロア 3　NISTEP の林（SPARC Japan 運営委員）です。クロスチェックでリジェクトしているとおっしゃっていましたが、もしご存じだったらどこで切っているのか教えていただけますか。あれはどこで切るのか実際使ってみるとものすごく悩ましいのです。単純にパーセンテージが何パーセントだからという基準ではなさそうですが、あまり手間を掛けているようにも思



（図 20）

えず、ご存じの範囲で教えていただければ。

●**引原**　知っている範囲だけですが、それも質問がありました。実際にやって、これはどうだろうかとなったときに、エディトリアルとは言いませんが、少し上のレベルで判断するというプロセスを経ています。それで、機械的にやるのではなく、本人に問い合わせるということもしています。現に、私自身が問い合わせられたことがあります。そこはセンシティブなので、まめにやっています。

●**フロア 3**　結構、丁寧にやっているということですね。

●**引原**　丁寧にやっていると思います。

●**フロア 4**　日本学術振興会の関係者です。私は物理系なので比較的 arXiv.org には慣れているのですが、arXiv.org はプレプリントサーバですから当然プレプリントなわけですね。それが最終的に投稿したジャーナルの何巻の何ページに出たという情報が arXiv.org にどのくらいのパーセンテージで反映されているかはご存じでしょうか。それは本人がやらないとできないようになっているのでしょうか。

●**引原**　新しいものはよく分からないのですが、古いものに関して言えば、本人がその情報を書いた場合に示されています。メタデータがここ 2 年ぐらいで大きく変更されました。以前はファーストオーサーだけがメタデータに入っていて、セカンド、サード、それ以降は出ていなかったのですが、それが今は変更されています。

そういう段階になったので、著者管理に関しても非常に明確になってきました。ですから今後は、resubmission も含めて、きちんとどこに出たかというものも合わせて出す方向で進めるのだろうという議論にはなっていますが、まだ確定ではないです。

●**フロア 4**　ドイツでは自動的に arXiv.org に上がるシステムになっていると初めてお聞きして、何かそういう方向に進むのではないかという気がしています。

●**引原**　それはエンカレッジしています。例えば、日本の学会はどうですかと聞かれて、科学技術振興機構などにお願いするのがいいのではないかという話は当然あるのですが、オープンアクセスの考え方と整合を取らないといけないので、国としてきちんと定めた流れの中でやるというのが重要だと思います。

ドイツの場合はまさにそれをやっていて、現在、Elsevier とオープンアクセスに関して闘っていますが、その状況の中でこれは重要ですし、これと同時にやりとりするための ILL を構築していくというのがドイツの考え方です。入れる側と出す側の両側をきちんとグリーンにしていこうというのが戦略としてあるのだと思います。そこをサポートできると思います。

第 2 回 SPARC Japan セミナー2017

「プレプリントとオープンアクセス」

学術情報共有とオープンアクセスの未来

Gregg Gordon
(Managing Director, SSRN)

講演要旨

オープンアクセスとデジタルリポジトリは研究の相互交流において重要な要素である。これまでのグリーン OA とゴールド OA に最近加わったブラック OA はシンプルに回答できない課題を呈示した。『共有』はこの 20 年以上に渡って明らかに変わってきた。演者は SSRN の概要を説明し、社会科学分野のみでスタートしてからその範囲が生命自然科学を含め 30 を超える分野へ広がってきたことを紹介する。本講演では現実に起きている出来事や、広範囲での学術情報共有が引き起こす問題、オープンアクセスと学術情報流通の未来への考えを述べる。

Gregg Gordon

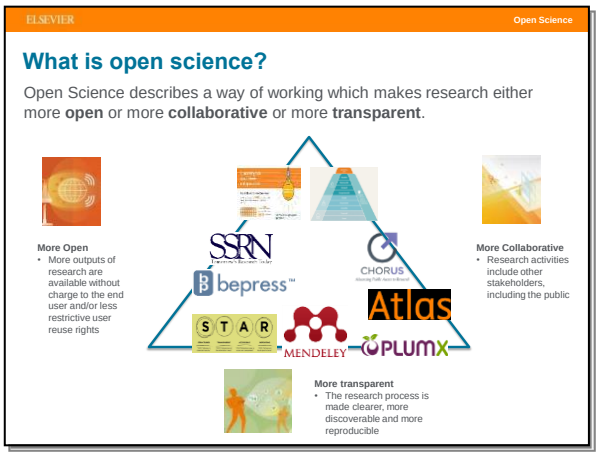
KPMGやテクノロジーやヘルスケア関連の企業に勤務後、1994年にSSRNを創始したMichael C. Jensenを補佐する。学術研究と、イノベーティブな研究をより早めるために必要となる変化について世界中で講演し、定期的に執筆している。最近の著書に、PLOS Biologyから出版されたThe Question of Data Integrity in Article-Level Metrics（共著）がある。

これから、Elsevier の土台であるオープンサイエンスと SSRN について、そして私たちがどんなふうに世の中を見ているかについて、お話しします。幾つかポイントを取り上げて、全体像を示したいと思います。

初めに、広い観点から Elsevier を見てみましょう。Elsevier は RELX グループの一部で、RELX はさまざまな企業を抱えていますが、どの企業も、情報の共有をベースにしています。RELX は、世界 24 カ国に 66 以上のオフィスを持つグローバル企業で、「The Lancet」「Cell」などの定評あるブランドを抱えています。出版業の変化に伴い、Elsevier も変化しています。オープンサイエンスに焦点を絞り、単なる出版社を超えたデータアナリティクス企業になりつつあります。

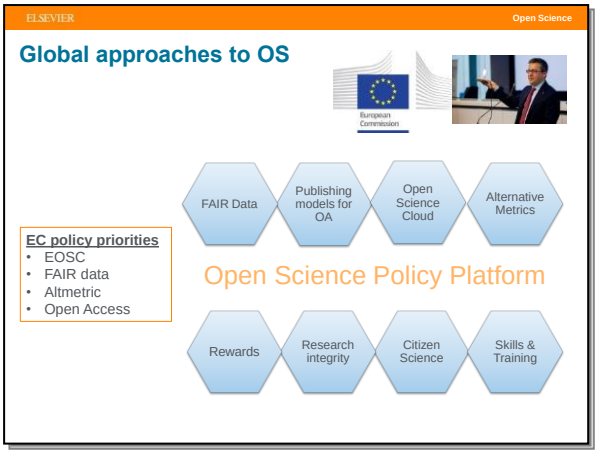
オープンサイエンス

オープンサイエンスとは、研究をもっとオープンに、もっと協力的に、もっと透明性のあるものにするために協力して作業する方法のことです。ご覧のスライドには、「オープン」、「協力」、「透明性」という異なるポイントに貢献する Elsevier の幾つかの要素が示され

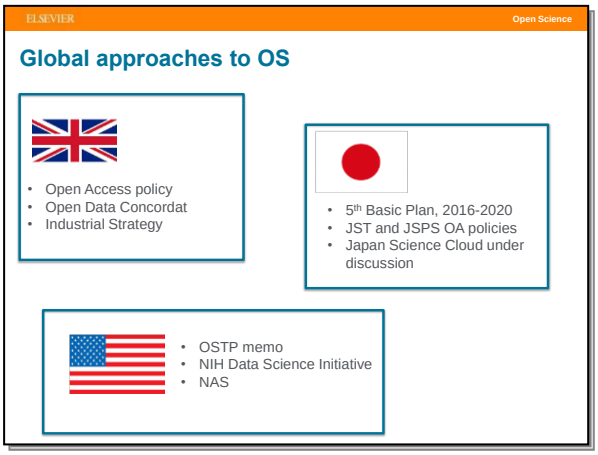


（図 1）

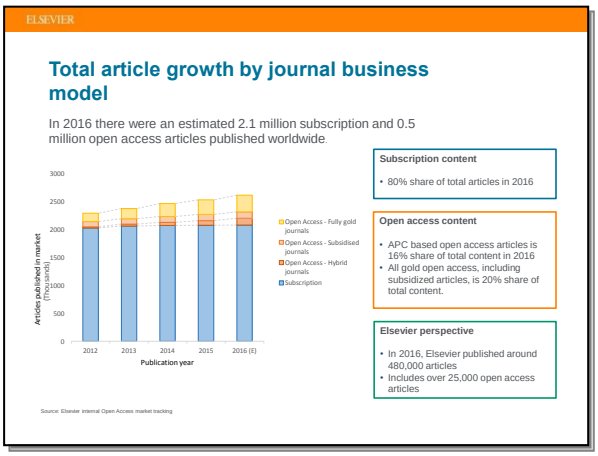
ています（図 1）。SPARC や ORCID など、世界中の組織が関わって、世界中でたくさんの活動が行われています。また、イギリス、日本、アメリカで、オープンサイエンスに向かう世界的なアプローチが取られています（図 2、図 3）。



（図 2）



（図 3）



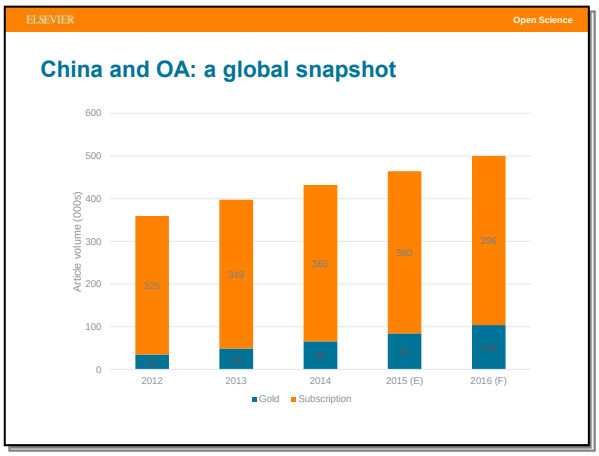
（図 4）

よりオープンに

まずはよりオープンになることについて話しましょう。Elsevier におけるオープンアクセスの定義はシンプルで、誰でも学術研究に無料で永続的にアクセスできるということです。このオープンアクセスには、簡単に言えばゴールドオープンアクセスとグリーンオープンアクセスがあります。このグラフは、いろいろなタイプ別のオープンアクセスの伸びを示しています（図 4）。しかしながら、論文掲載の大半は依然として購読ベースです。この地図は、世界でのグリーンオープンアクセスの利用状況を示したものです（図 5）。これらの国々の多くにはゴールドオープンアクセスもありますが、グリーンオープンアクセスの方が世界中で利用でき、普及しています。

中国を例に見てみましょう。中国はアジア諸国の中で最も成長の著しい国です（図 6）。しかしながら、

（図 5）



（図 6）

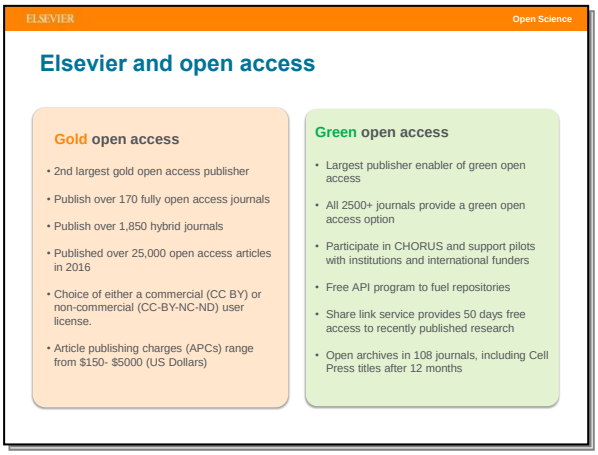
成長は依然として限定的です。私は、これは私たち全員にとって、より共有を進めるチャンスだと捉えています。図 7 はゴールドとグリーンそれぞれの、Elsevier のオープンアクセスに関するデータの一部です。

もっとシェアできるようにするための方法の一つは、Elsevier のアプリケーション・プログラミング・インターフェース（API）を活用して、メタデータとコンテンツのいずれも機関リポジトリに入れることです（図 8）。Elsevier は CHORUS にも取り組んでいます。現在アメリカの諸機関と実験をしていて、日本とオーストラリアでも試験サービスを行っていますが、もっと幅広い機関リポジトリにコンテンツを提供するというものです。もっとシェアするためのもう一つの方法は、データの共有を始めることです。データはたいてい複数の研究者が利用しますから、シェアすればするほど、研究者にとってメリットが大きくなります（図 9）。

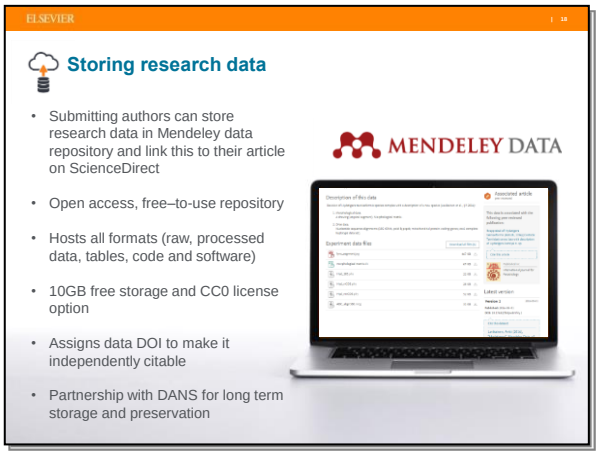
データを論文にリンクさせると、特に特定分野のデータリポジトリに保管されている場合、とても役に立ちます。現在では、SSRN 内でデータを共有したり、多数の分野でデータ論文を発表したりすることが可能です。最新のパイロット試験ではオープンピアレビューの共有を開始しました。一般的なピアレビューの問題の一つは、査読者が匿名で、その功績が認知されていないことです。オープンにして査読者の認知度を上げることによって、もっと多くの査読者が査読をするようになると私たちは確信しています（図 10）。

より協力的に

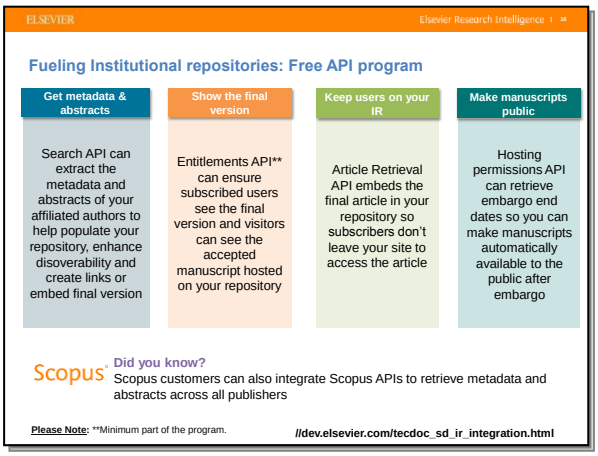
協力についてお話ししましょう。私は中国科学院の人と話をしたのですが、中国科学院はたくさんの国際機関と協力しています（図 11）。SSRN もその協力機



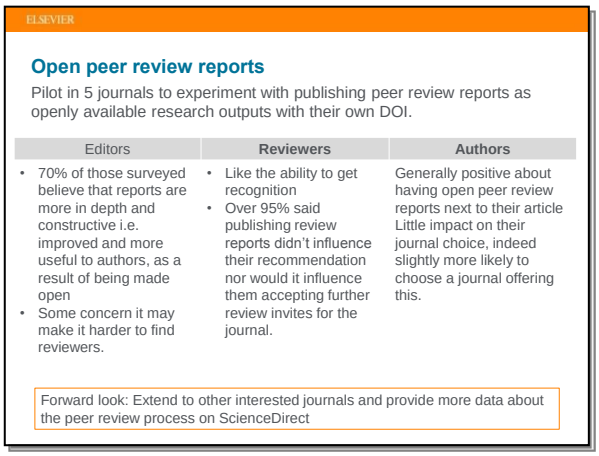
（図 7）



（図 9）



（図 8）

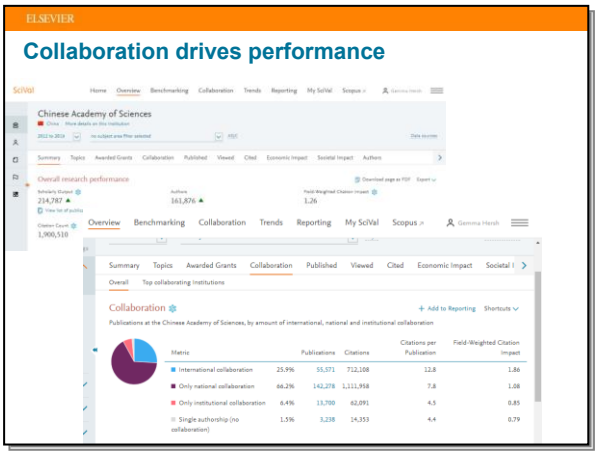


（図 10）

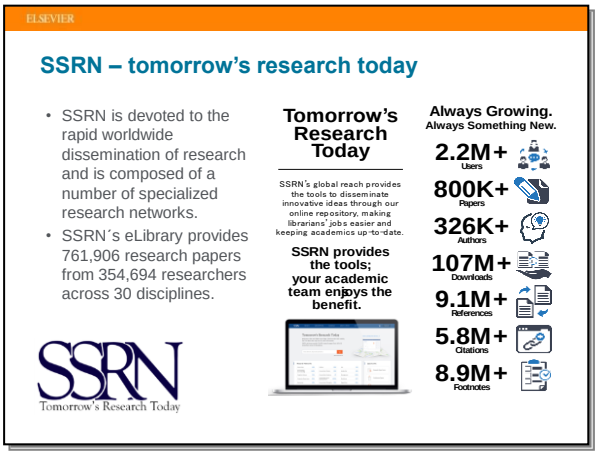
関の一つになっていて、SSRN の利用者、論文、著者の数からそのことが分かります（図 12)。Mendeley はもう一つの協力例で、参考文献やその他の文献を初期段階で共有するものです。他者との協力という点では、社会的影響を評価する Atlas Award もあります。AudioSlides は、他分野の専門家や研究者に動画や 5 分間のプレゼンテーションを提供します。図 13 は、エルゼビアの最新の報告書「Gender in the Global Research Landscape（世界の研究環境におけるジェンダー)」です。SSRN が今週、初期段階の研究共有に向けて Women's & Gender Studies Research Network を立ち上げたことが示されています。

より透明に

次に透明性ですが、私たちも他の多くの出版社にならって、科学コミュニティとともに研究を透明性のあるオープンなものにするためのデータガイドラインを採用しています。また、私たちはデータベースの Scopus で、検索をよりしやすくしました。とても重要なツールである Scopus が、もっと幅広く使えるように改善されたのです。DataSearch は面白いツールで、メタデータを検索するだけでなく、自分でデータそのものを検索できるので、全体をダウンロードする前にそのデータセットに関心を持てるかどうか、判断することができます。また、私たちはデータの引用を増やすために、FORCE 11 とも協働しています。データの引用は、論文の引用とは種類が異なりますが、私たちはデータセットの作成者が認識されることも重要だと考えています。そして、ネガティブな結果となった情報の報告も強化しました。そうすれば他の人たちの研究から学べるからです。さらに、データだけを掲載する専門ジャーナルもあります。最初にお伝えしたとおり、Elsevier はデータインフォマティクス企業になりつつあります。



（図 11）



（図 12）

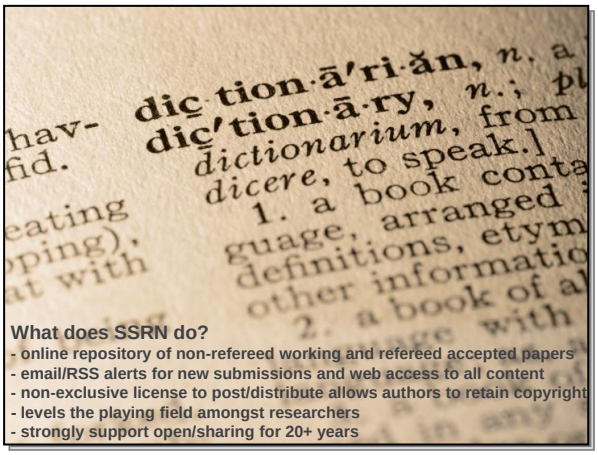
（図 13）

SSRN について

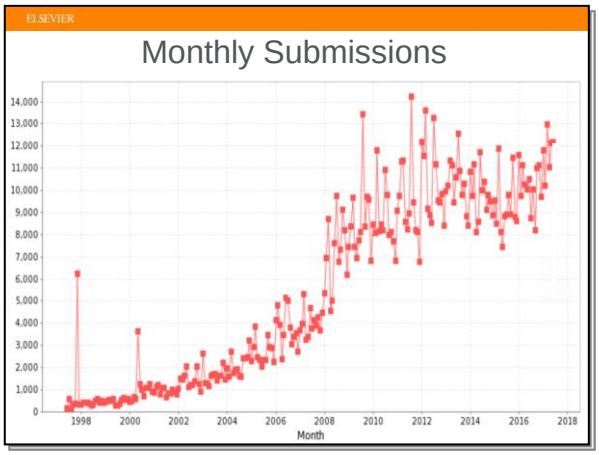
SSRN は 22 年ほど前にスタートしました。図 14 がその基本情報です。これらのポイントのうち一番重要なのは四つ目の、研究者に公平な競争の場を提供するということだと思います。ハーバード大学でも、私が住むニューヨーク州ロチェスターのモンロー・コミュニティ・カレッジでも関係ないということです。出身の国や地域、どの言語で研究を行うかも関係ありません。

SSRN には約 35 万人の著者がいます。図 15 のグラフで急に飛び出ている部分は、他の組織が SSRN に参加したときです。図 16 のグラフは論文投稿の数です。同様に飛び出ている部分があります。また、ここ数年で大幅に増加しました。図 17 は重要なスライドだと思います。青の点は 1 日の投稿数で、赤の点は 1 日の修正数です。2007 年 11 月に、私たちは SSRN ライブラリの論文の修正やアップデートを認めるようになりました。このグラフは 2012 年 4 月で終わっています。2012 年 4 月のこの日には、新規投稿より修正の方がたくさんありました。これは、SSRN の発展において重要な日です。SSRN が生きている、呼吸している情報の集まりであることを意味しているからです。私たちは、時間とともに研究が発展し、より良くなっていくことを望んでおり、早期の共有によってその速度は早くなります。

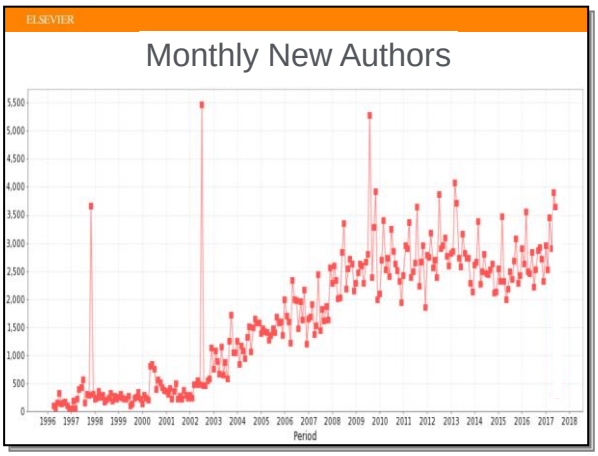
初期段階の研究にアクセスできれば、利用できる、自身に関係のある研究が増えます。研究者のアイデアや草稿、ワーキングペーパー、会議録、プレプリントを、最終的に発表された論文と同じように利用することができます。これまでは一つのバージョンしかなかったものが、いまや三つのバージョンがあります（図 18）。さらに、発表後のバージョンも追加し、研究を



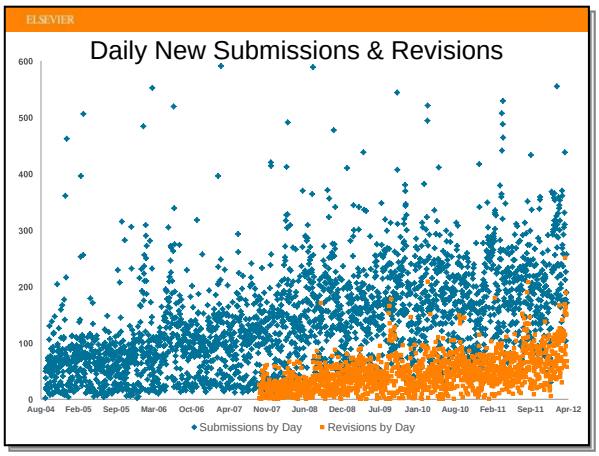
（図 14）



（図 16）



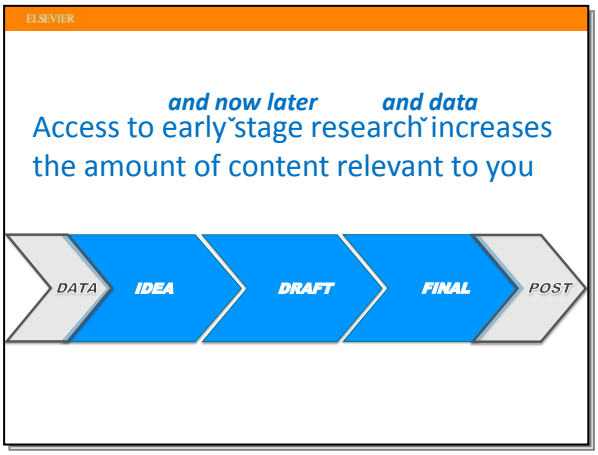
（図 15）



（図 17）

アップデートできるようにしました。データも追加すると、一つの研究論文から最高で五つのバージョンが出せるようになりました（図 19)。これらのバージョンの管理は複雑です。図 20 は SSRN 内の法学研究分野でのその複雑さの一例を示したものです。

（図 18）



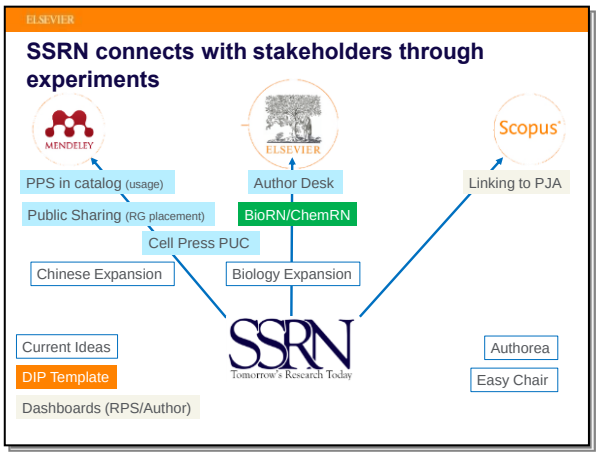
（図 19）

（図 20）

SSRN は、社会科学でスタートして、人文科学、生物学、化学が追加され、今週には女性・ジェンダー研究が追加されました。来月には工学を追加します。私たちのスローガン、マントラは、「Tomorrow's Research Today（今日、明日の研究を）」です。2014 年の「Journal of Financial Economics」に発表された論文は、一つを除いて全て、3 年半前には SSRN の eLibrary データベースで発表されていました。世の中に目を向けると、パートナーなしに私たちの運命を全うすることはできないということが分かったので、Elsevier と提携することにしました。それによってギャップを埋めることができたと考えています（図 21)。今、私たちは「Tomorrow's Research Today」から出発して、研究者をつなぎ合わせ、アイデアを展開させ、初期の研究をシェアする手助けをしているのです。

私が考えるイノベーションのシンプルな定義は、よ

（図 21）



（図 22）

り幅広い、より深いものに触れることで、新しいものを生み出すということです。私たちは、SSRN が最も幅広く、最も深い一連の研究を提示できるようにしたいと考えています。これは、研究パフォーマンスを改善したり、研究者が素晴らしい、新しく革新的な研究をもっと素早くできるよう支援したりするためです。私たちはギャップを埋めてきた一方で、実験やパイロット試験で実際にいろいろなピースをつなぎ合わせて、研究をもっと先に進めようとしているのです（図 22)。私は最近、SSRNblog.com のブログにオープンアクセスウィークを祝う投稿をして、過去 15 年間でオープンアクセスがどのように発展してきたか論じました。オープンアクセスの良い面、オープンアクセスのためにもっとうまくやらなければならないことについても論じています。ぜひご覧ください。

中国、韓国、日本での SSRN 利用統計

私はこの 1 週間半で北京、ソウルを訪ね、東京にやってきました。私たちは、いろいろな学校でさまざまな研究者が SSRN を活用する方法を見てきました。この数字は、中国（図 23)、韓国（図 24)、日本（図 25）での利用に関する統計を示したものです。研究者に、母国語で執筆し、研究し、共有できることを知らせるために、もっとやらなければならないことがあるということが分かりました。また、私たちは研究者に、SSRN は機関リポジトリ、分野ごとのリポジトリ、国別リポジトリとしても使えるということを知らせる必要もあります。

●フロア 1　バイオサイエンスデータベースセンターの職員です。中国と韓国と日本のユーザー数の比較で、日本のユーザー数がとても少ないように見られました。その理由は何だと分析されていますか。

●Gordon　他の講演者の方々とのランチのときにその話題が出ました。中国の方がコンテンツは多いですが、SSRN を使いはじめたのは日本の方が先です。日本の研究は他の 2 カ国に比べて、早くからもっと洗練されていたと思います。私は理解を深め学ぶためにここに来ているわけですが、私たちは世代交代の時期に差し掛かって

ELSEVIER

Researchers from China SSRN share their research on SSRN

Top 10 Business Schools
(Total Downloads)

School / Department	Authors	Papers	Downloads
The University of Hong Kong - Faculty of Business and Economics :: China	96	750	201,625
Tsinghua University - School of Economics & Management :: China	105	470	108,182
Peking University - Guanghua School of Management :: China	89	360	92,706
Cheung Kong Graduate School of Business :: China	21	163	84,501
China Europe International Business School (CEIBS) :: China	46	268	80,546
University of International Business and Economics (UIBE) :: China	102	245	26,800
Fudan University - School of Management :: China	50	137	22,527
Peking University - Shenzhen Graduate School of Business :: China	27	117	20,118
Shanghai Jiao Tong University (SJTU) - Antai College of Economics and Management :: China	49	123	18,730
Peking University - HSBC School of Business :: China	13	65	18,535

SSRN　Source: SSRN Website　- 58 -

（図 23）

ELSEVIER

Researchers from Korea SSRN share their research on SSRN

Top 10 Business Schools
(Total Downloads)

School / Department	Authors	Papers	Downloads
Korea University Business School (KUBS) :: Korea	53	221	76,540
Seoul National University - College of Business Administration :: Korea, Republic Of (south Korea)	30	150	57,650
KDI School of Public Policy and Management :: Korea, Republic Of (south Korea)	57	201	30,297
KAIST Business School :: Korea, Republic Of (south Korea)	48	128	22,249
Yonsei University, Seoul Campus, College of Business and Economics :: Korea	21	122	14,444
Korea Advanced Institute of Science and Technology (KAIST) - College of Business :: Korea, Republic Of (south Korea)	47	112	11,653
Sungkyunkwan University - School of Business Administration :: Korea, Republic Of (south Korea)	12	35	8,722
SolBridge International School of Business :: Korea, Republic Of (south Korea)	9	23	5,742
Gachon University - College of Business and Economics :: Korea	17	64	4,749
Hankuk University of Foreign Studies - College of Economics and Business :: Korea	8	19	4,018

SSRN　Source: SSRN Website　- 59 -

（図 24）

ELSEVIER

Researchers from Japan SSRN share their research on SSRN

Top 10 Business Schools
(Total Downloads)

School / Department	Authors	Papers	Downloads
Tokyo University of Science - School of Management :: Japan	4	30	7,911
Kobe University - Research Institute for Economics & Business Administration :: Japan	7	43	5,921
Kobe University - Graduate School of Business Administration :: Japan	21	42	4,894
Hitotsubashi University - Graduate School of Commerce and Management :: Japan	16	35	3,490
Hokkaido University - Graduate School of Economics & Business Administration :: Japan	17	57	3,331
Aoyama Gakuin University - Graduate School of International Management :: Japan	3	13	2,531
Hosei University - Graduate School of Business Administration :: Japan	1	13	2,432
University of Tsukuba - Graduate School of Business Sciences :: Japan	3	8	2,371
Keio University - Faculty of Business and Commerce :: Japan	10	30	2,143
Keio University - Graduate School of Business Administration :: Japan	4	13	1,986

SSRN　Source: SSRN Website　- 60 -

（図 25）

いるのではないかと思います。以前はより先進的な日本人研究者の多くが SSRN を使っていました。今は、もっと若く、頭角を現している研究者が SSRN を使うようになってきています。私は、彼らができるだけ早く成功を収められるよう支援するためにやってきました。

●フロア 2　横浜国立大学の深貝と申します。私は SPARC Japan の運営委員の一人です。商業ベースの仕組みと、arXiv.org のように学者・研究者が始めた仕組みはなじまないところがあるかもしれないということ、これが今後どうなるのかは大切だと思います。

今お話しになったように、世代交代があれば、SSRN、あるいはパブリッシャーが提供するものを使う人の比率が増えるのは確かだと思います。それは便利なものになじむ世代が増えるということですが、それで学術がうまくいくかというと別の問題です。産業革命から現代に至るまで、いろいろな技術革新がありましたが、単に新しいものを使うことが大事なのではないのです。大事なのは、そこからどういう新しい文化をつくれるかということです。今、生活において、世の中の在り方において、新しい技術の意味が問われています。学術においてもどうかということをわれわれは考えなくてはいけません。

オープンサイエンスの中でのツールは誰が提供しているのか、自分たちでつくっているのか。商業ベースで提供していただけるのは確かに便利であって、ありがたいですが、その中でどういう研究が育ち、その知識がこの世の中においてどういう貢献ができるかということは、それはそれで別に考えなければいけないのではないでしょうか。

●Gordon　とてもいい質問ですね。少し話を絞ってお答えしましょう。私は、単なる商業対学術という見方はしていません。私たちは、少なくとも各国の政府と資金援助団体を巻き込んでいかなければなりません。

私の友人に Simeon Warner と Paul Ginsparg がいます。Simeon は、arXiv が持続可能なビジネスモデルを模索していたとき、ロチェスターにやってきました。何に、どう課金できるか、そして一番重要なこととして、どう自ら資金を調達するかについて、私たちはたくさんのシナリオを検討しました。商業面の問題は、bioRxiv や Open Science Framework などを見てみるともっと複雑です。bioRxiv は現在、フェイスブックを創業した Zuckerberg の Chan Zuckerberg Initiative（CZI）の資金提供を受けています。Open Science Framework は、億万長者のヘッジファンドマネージャーが運営する Laura and John Arnold Foundation の資金提供を受けています。

私たちの選択肢の一つは、政府です。政府は有権者、党員、特別利益団体に対してその活動の正当性を説明しなければなりません。他の選択肢としては、bioRxiv や Open Science Framework のように、億万長者のビジネスマンが非営利の資金提供を行い、このような手段を通じて意思決定をしてくれるということがあります。三つ目の選択肢は、営利組織です。彼らは価値を提供できなければ、顧客を失ってしまいます。私は、このうちのどれが本質的に良いとも悪いとも思いません。どれにも欠点はあります。それぞれに欠点があるということを認めない人にはいらいらさせられてしまいます。

私は、SSRN が Elsevier に加わったことにワクワクしています。Elsevier は巨大な営利企業で、良いことばかりではありませんでした。この共有というゲームは非常に高くつくものになっていて、自力では成功できない、と私たちは分かっていました。幸運なことに、自分たちがしていることの可能性を信じ、全面的にサポートしてくれる大きな組織と提携することができました。ふさわしいパートナーがいなければ、生物学や化学、女性・ジェンダー研究、工学の全てを今年追加することはできなかったでしょう。私としては、心変わりして、資金提供をやめてしまう組織に頼るよりも、みんなが去ってしまうような状況でも価値を提供し続けなければならない組織を信頼しています。私たちはただ、SSRN はしっかりと管理されており、情報のセ

キュリティが確保されているということを保証しなければなりません。それが、ずっと長期にわたって仕事を続けられるようにするために、私が SSRN の中で最善を尽くそうと努力してきたことです。でも、私はこの週末東京に来ていいお天気を楽しんでいますけれど、それと同じでどこにも保証はないのです。

●フロア 3　NISTEP の林（SPARC Japan 運営委員）です。2013 年に、Fiesole Retreat という出版社を中心とした非公式会合があり、セッションは違いましたが、Gregg さんと一緒に講演してディスカッションする機会がありました。そのときのエピソードを交えてシンプルな質問があります。

2011 年に SPARC Japan セミナーで Mendeley の CEO を呼んだ際に、「この先、どうするの。もしかして Elsevier が買ってしまったらどうしよう？」と話したところ、彼は明らかに反応が変わり、その後、Mendeley は Elsevier に買収されました。

2013 年に私は Gregg さんに「ビジネスサステナビリティはどうされるのですか。商業的な資本が入る可能性はありますか？」と質問しました。そのときは、SSRN としてブランドを大事にしていきたいというお話をされましたが、今は Elsevier の傘下にあります。

私は非常にニュートラルに、今まさにおっしゃっていたように誰も将来が見通せない中、いろいろチャレンジをしていく上で、パトロンが誰であるかが非常に大事であるということを理解しています。

その上で質問が二つあります。一つ目は、SSRN が Elsevier の中でどうサステナブルかということです。二つ目は、今の立場から今の arXiv.org のビジネスサステナビリティはどのように見えるかということです。Taxation model と呼ばれることもありますが、会員制で、会員からお金を集めて、その代わり公共にフリーに出すというモデルについて、ご自身の見通し、パースペクティブを教えてください。

●Gordon　まず、Mendeley と SSRN にはいろいろな選択肢があったことを言っておくべきだと思います。正直なところ、資金的にはもっといいオファーがあったのですが、SSRN の発展にとっていいものではありませんでした。この部屋の中で、20 年以上同じ仕事をしている人は何人いますか？　ほんの少しだけですね。それだけ長い間何かができるということは、大好きだからやっているということです。いろいろな問題、いろいろなフラストレーションにもかかわらず、続けていきたいわけです。私たちは、できる限り SSRN を今のまま続けていくことを保証できる取引を構築しました。

客観的に見てみると、Elsevier もその他の全ての出版社も変化していることが分かります。愚かな人たちではありません。秘密の部屋で意思決定をしているわけではありません。あらゆることがオープンなところで起こっています。みんながいろいろなことをしていて、私たちの誰も、20 年後、10 年後、あるいはたった 5 年後でさえ、未来がどうなっているか分かりません。私は自分のチームや、Elsevier 内の他の人たちにも、「今はクレイジーな時代だ。それを受け入れるか、そうでないなら違うことをしなさい。でも、重要なのは間違った答えは存在しないということだ。このことを毎日理解しておく必要がある」と言っています。

最初のご質問に戻りますと、私は学会やその他の会員制組織と過ごす時間が多いのですが、皆さん会員の維持に苦労しています。会計学を学んだ身としては、このようなビジネスモデルは採用しません。ですが、毎年何十万マイルも飛び回って、あらゆる場所でここにいる皆さんのように賢い人たちの話を聞いている身ですから、私が間違っている可能性も大いにあります。もっと大切なことは、このような方々の成功を、私は心から願っているということです。

第 2 回 SPARC Japan セミナー2017

「プレプリントとオープンアクセス」

化学分野におけるプレプリントの位置付け・課題等について

生長 幸之助

（東京大学大学院薬学系研究科/化学ポータルサイト Chem-Station 副代表）

講演要旨

現行の査読システムに諸々の問題点が見えつつある状況下において、迅速な研究情報共有プラットフォームとしてのプレプリント投稿が各分野で台頭しつつある。化学分野でも、アメリカ化学会主導で ChemRXiv というプレプリントサーバの運営が先日より開始され、注目を集めている。本講演では、プレプリント投稿が今後化学分野でどのような普及・活用・発展を遂げていくのかについて、事例を参照しつつ議論してみる。演者は大学にて教員・研究者として働く傍ら、化学分野の国内最大二次情報プラットフォーム Chem-Station の運営にも長年関わっている。研究者とメディア運営の二足のわらじを履く立場からの視点も併せて述べてみたい。

生長 幸之助

1980年徳島生まれ。2003年 東京大学薬学部 卒業。2008年 東京大学大学院薬学系研究科 博士課程修了（指導教員：柴崎 正勝）。 同年 博士（薬学）。 2008-2010年 カリフォルニア大学ロサンゼルス校 化学/生物化学科（Omar M. Yaghi研究室）にJSPS海外特別研究員として赴任。2010-2016年 東京大学大学院薬学系研究科 助教。2016年より現職（講師）。化学ポータルサイト Chem-Station 副代表（2002-）、ERATO金井分子触媒生命プロジェクト研究総括補佐（2015-2017）を兼任。

私の名前は生長（おいさき）と読みます。大学で教員を務めていますが、本業は実験化学の研究者です。バイアルやフラスコに化学物質を混ぜてものをつくるということをやってきました。自分なりの立場からこのプレプリントについての見方や意見、将来どうなっていくかということをお話しできればと思います。

1.自己紹介

私の専門は薬学の中でも有機合成化学という分野です。もっと狭い話をすると、触媒反応開発という分野になります。一番皆さんになじみのある研究者の名前を挙げると、2010 年にノーベル賞を取られた、クロスカップリングの鈴木先生や根岸先生の分野に一番近いです。この薬という複雑な構造を持った有機化合物に対して、もっと良いものにしたい、どうやったら効く分子になるかということを考えていろいろ加工していきます。

これを一度にやろうとすると大変なのですが、魔法の粉といいますか、特殊な化学反応を進行させるような触媒を自分たちの手でテーラーメイドし、ショートカット反応を起こして薬を効率的に供給したり、今までアクセスできなかった構造の化合物にアクセスできるようにしたりしようという発想で研究をしています。この触媒を開発するというのが私の主な研究テーマの一つです。長い目で見ると、薬開発を促進させるというビジョンを持ってやっています。

図 1 は、触媒開発の一環で出てきたごく最近の成果で、一番分かりやすいものです。今は小さな有機化合物だけではなく、抗体なども薬になっています。かなり大きなサイズの分子も薬になっています。タンパク

質なども薬になるという時代になっているのですが、これを人間の力で加工することができれば非常に面白いのではないか、そういう化学反応を自分たちの手で開発しようというテーマにも取り組んでいます。

最近、触媒の基になる試薬がタンパク質のアミノ酸であるトリプトファン残基にくっついて、トリプトファン選択的に化学反応を起こすというものを開発しました。これを使って抗体薬を化学修飾して、もっと強い効き目を持つ抗体薬に変えるという研究にも現在取り組んでいます。

このように化学反応ベースで新しい薬に道を付けようという考え方で研究をしている人間です。

2.Chem-Station について

2-1.運営体制

ここまでが研究の話なのですが、これ以上に私は長年 Chem-Station という化学のサイトの運営に携わっています。

Chem-Station は図 2 のような見た目のウェブサイトです。「ケムステ」と検索を打てば出てきます。これは一体どういう化学サイトかというと、日本最大のアクセス数を稼いでいる、化学の情報を集約してまとめて発信する、いわばポータルサイトと呼ばれるものです。

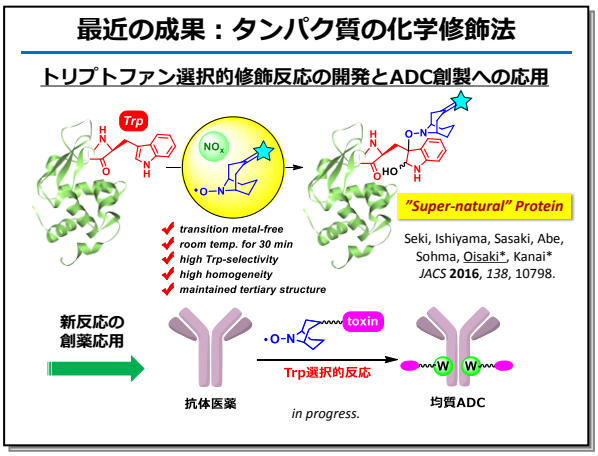
設立は 2000 年で、私は 2 年目ぐらいから、約 16 年研究室に入る前からずっと運営に関わっています。このサイトは「化学とウェブの融合」という大きな目的を掲げていて、将来的に化学とウェブを融合させることができないかということを考えて日々やっております。具体的には、ウェブにたくさん存在している化学に関する情報を分かりやすい形でまとめて提供するということをほぼ人力でやっている形になります。

これだけ長く続けているので、知名度も最近すごく上がってきて、アクセス数が 80 万～240 万 pv/月となっています。

公開ページ数も、これは蓄積の産物だと思うのですが、5,000 ページを超えていて、ボランティアも含めていろいろなスタッフの方にご協力いただいています。パーマネントのスタッフはおらず、本業がある中で協力いただいています。

運営も、ある程度持続可能なものにしたいため、アクセス数があるページなので、スポンサーから広告を出してもらって、サーバー代、記事のお金に換えていくというようなやり方をしています。化学系の会社と長いお付き合いがあります。

私は副代表という位置付けで、代表は早稲田大学理工学術院准教授の山口潤一郎さんです。彼がサイトを立ち上げて、私が後で加入しました。二人が一番古株なので、代表、副代表という関係でずっとやっています。最初はスタッフの数は 2 人ぐらいしかいなかったのですが、年を追うにつれて知名度が上がって、やりたいという人も募れて、今ではスタッフの数が 100 名を超えています。他にもサポートしてくださっている方がたくさんいます。



（図 1）

（図 2）

このサイトを運営しているのはどういう方かというと、基本的には皆さん本業があって、ボランティアに近い形でやってくださっている方が多いです（図 3)。これはかなり特徴的かと思うのですが、実際にスタッフがどういう学歴を持った方かということをまとめ直してみると、博士課程の学生か、博士号を既に持った方が約 7 割を占めて、そういう人たちが化学のウェブ上の記事や専門的な文献などを読んでまとめて日本語で記事をつくってオープンにするという作業に関わっています。スタッフの目がしっかりしているということで、信頼性も担保されています。

もともと研究的な視野から出てきているサイトなので、研究者に役立つ情報を分かりやすく提供しようという考え方がかなりあり、必然的にスタッフの本業も研究であることが多いです。端的に、ケムステをつくっているのは研究者であると言うことができます。

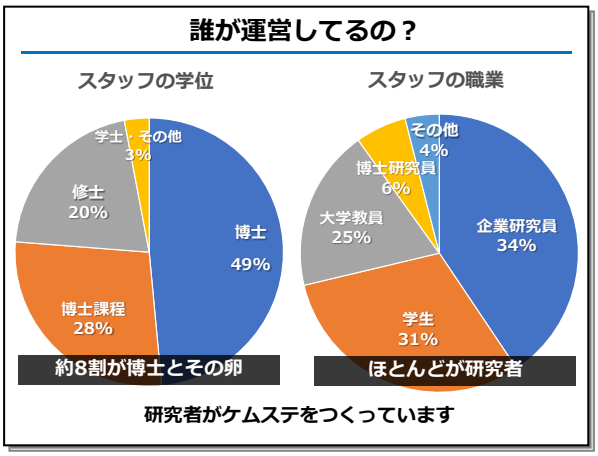
2-2.発信内容

研究者がどういう情報を発信しているのかというと、実は研究者専門というわけではなく、一般市民にも分かるぐらいのところまでカバーして、スタッフの書きたいようにやってもらっています（図 4)。研究者ならではの視点からその情報を読み解いて、分かりやすく文章にして発信するブログを書いたり、ニュースの形で紹介したりします。ニュース自体は一般の新聞が取り上げているものなのですが、それに例えば化学者ならではのコメントを付けて、解釈して出してあげることで付加価値が出ています。

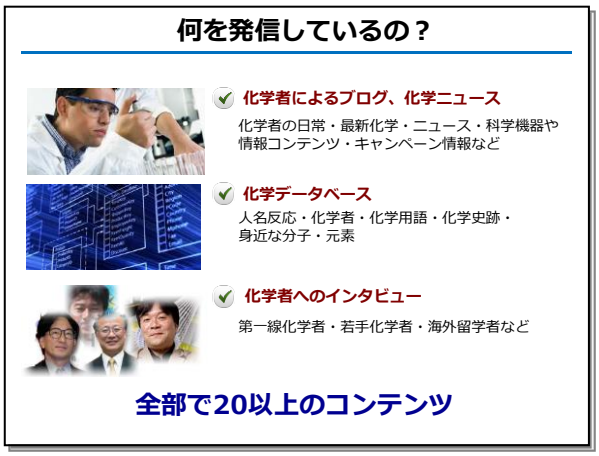
こちらも特徴的ですが、化学者ならではの必要な情報を整理してまとめ上げたデータベースをつくっています。かなりアクセス数があります。また、化学者がつくっているサイトということで、有名な化学者はどういうことを考えて研究に取り組んでいるのか、毎日どうやって過ごしているのかという生の声をお届けするようなコンテンツもあります。サイエンスコミュニケーションの走りのようなことをずっとやってきて、今そのような形で認知されているというサイトです。今は大変バラエティに富んだコンテンツが提供されています。

ある化学情報を、日本人に分かりやすいように日本語で発信するということに一番ニーズがあります。しかし、外国にも届けないと、日本の化学がいくら良いものでも、全体として知名度が上がらないですし、外国からクレームが付いたときに反論できる土壌がない、そういうことが実際に別の件であったので、そこに問題意識を向けました。基本的には日本語のサイトを英語化して、中国人スタッフを抱えて中国語に翻訳して、二つのサイトを並走させながら日本の化学というものを世界に発信していくということもしています（図 5)。

また、サイエンスコミュニケーションの一環だと思いますが、化学が一番注目を集める時期はノーベル賞受賞の時期かと思います。それに合わせて、Chem-Station が旗を振って、誰がノーベル化学賞を取るのか



（図 3）



（図 4）

という予想投票企画を Facebook の機能を使って行っています。今年は 500 票ぐらいアンケートが集まりました。Facebook を使っている人が気軽にワンクリックで誰がノーベル化学賞を取るかいうことを投票できるような企画です。当たったら賞金が出ます。

それと並行して、ノーベル賞が発表されたら、どういう人たちがノーベル賞をどういう業績で取っているのかということを日本語の記事にしています。

普通の日本のメディアだと、日本人が取らない限りは情報がなかなか出てこないのです。私たちはそれはどうかと思っていて、ノーベル賞を取った素晴らしい業績なのだから、ぜひみんなにも知ってほしいということで、今年はクライオ電子顕微鏡について、自分の勉強も兼ねてきちんと調べて、日本の記事として書いています。それが巡り巡って化学を皆さんに知ってもらうような活動になっているのではないかと信じて、毎年やっています。

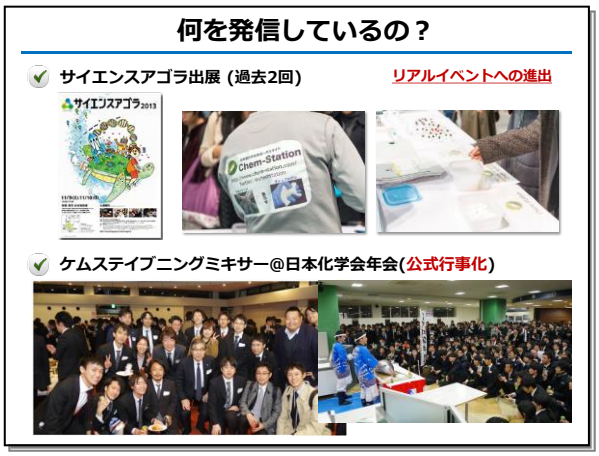
基本的にはそのようにウェブベースでの活動をしているのですが、リアルの活動にも少しずつ進出しはじめています。図 6 の上は、科学技術振興機構が開催している、「サイエンスアゴラ」というサイエンスを土台にしたお子様向けのお祭りです。それに Chem-Station のブランドで参加し、化学実験のブースをつくって、化学は面白いということをお子さんやそのご両親に体験していただくということをスタッフ主導でやっています。

また、このケムステというサイトは若い人がかなり見てくれていることが特徴です。30 代以下が読者として多く、それに日本化学会が目を付けました。日本化学会年会は毎年行われている学会で、懇親会も開いていますが、若者向けがない、開催してもおじいさん、おばあさんしか来ないという実情があったようで、何とか若者を引き留めたいということを、当時会長だった玉尾皓平先生が考えたのだと思います。そこでうちの代表に話を持ってきて、ぜひケムステ主導で若者の懇親会をつくってほしいということで実際に実現したのが「ケムステイブニングミキサー」というものです。下の写真は今年 3 月の懇親会の様子で、このときは 500 人ぐらい集まりました。

なぜこんなに集まったかというと、マグロの解体ショーを代表が自腹を切って企画したからということがあります。これを見たいがためにいろいろな学生たちが来て、大盛況になっていました。これは代表の企画力がすごかったという一例かと思います。

というわけで、Chem-Station のようなサイトは日本にはほとんどありません。本当に唯一無二と言ってもいいぐらいのレベルでやっているとわれわれは自負しています。その中でグローバルな発信も視野に入れながら、みんな本業がある中でやってくれているので、そこまで進むかどうかという問題はあるのですが、できれば世界一の化学メディアを目指していきたいです。社会に化学の情報を行き渡らせて、化学の好きな人を増やしていきたいというビジョンを持って続けております。

（図 5）



（図 6）

3.化学とプレプリントサーバー

3-1.化学系プリントサーバーChemRxiv

ここから今日の本題に入ります。ケムステという活動をしていて、プレプリントサーバーというものに個人的に興味を持ち、まとめて紹介記事を書きました。これがプレプリントサーバーについてここで話すことになったきっかけです。

最初に書いたのは、化学系プリントサーバーChemRxiv の設立が決まりましたというイントロ記事です（図 7）。これはまだできていない段階で取り上げて、化学のみんなはプレプリントサーバーを多分知らないだろうなと思ったので、プレプリントサーバーとはこういうものなのですよという追加の記事を付けました。

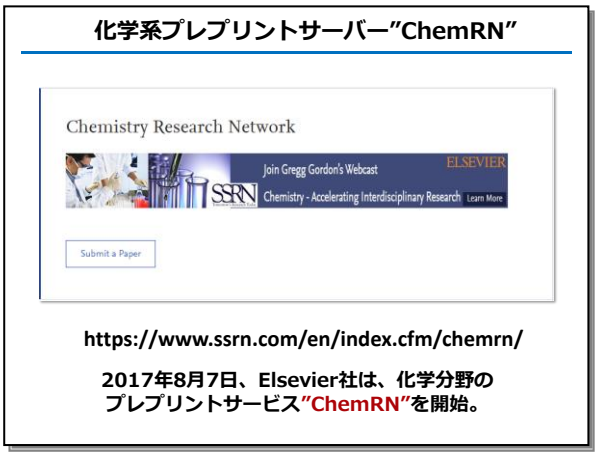
実際にベータ版ができるということが分かり、やりはじめましたということをニュース記事にして、ぱらぱらと見てみた感じの注意点などもまとめたブログ記事を研究者向けに出しました。この ChemRxiv が化学系のプレプリントサーバーとして米国化学会の手によって設立されて運用が始まったのが 2017 年 8 月の話です（図 8）。

先ほどの Gordon さんのお話に出てきたような、SSRN 主導で化学分野のプレプリントを集めるということも今まさに始まらんとしています（図 9）。こちらも 2017 年 8 月にアナウンスが出たところです。

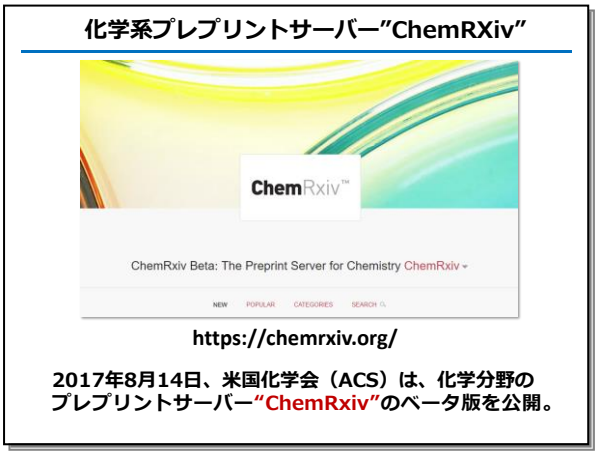
SSRN のことは横に置いておいて、ChemRxiv に話を絞っていきたいと思います。やはり始まったばかりの取り組みで、活発に ChemRxiv がコミュニティで使われているのかと言われると、まだまだこれからという実感はあります。

図 10 は ChemRxiv のサイトから引用したスタッツです。ChemRxiv の 2017 年 8 月 10 日から 10 月 15 日ま

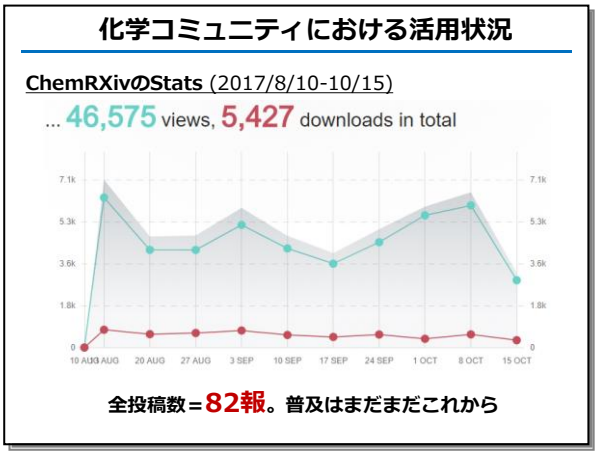
（図 7）



（図 9）



（図 8）



（図 10）

でのアクセス数は 46,575、ダウンロード数は 5,472 です。この短い 2 カ月の期間で、私が確認した最近の論文投稿数は全部で 82 報で、arXiv.org の 11 万報とは雲泥の差です。これはできたばかりなので仕方がない、まだまだこれからだとご理解いただくしかないと思います。

3-2.プレプリント投稿のメリット

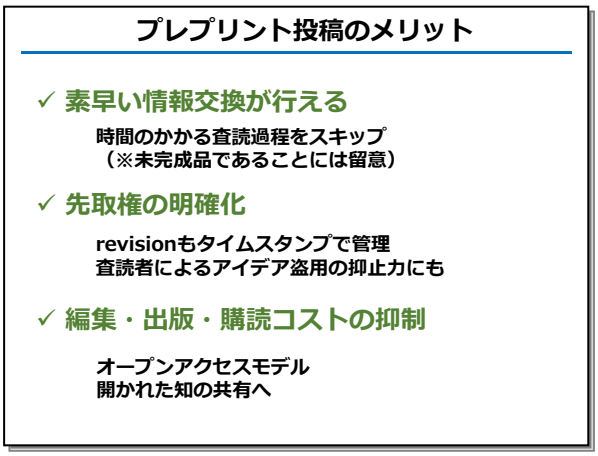
プレプリント投稿の大きなメリット（図 11）は、とにかく査読過程がスキップできて、素早い情報交換が行えるということです。ただ、査読を終えていないので考えて使ってくださいということです。また、誰が最初にその仕事をしたかということが査読を経ている間にうやむやになっていかないということが大きなポイントかと思います。査読中に、アイデアの盗用ということが問題になると思うのですが、ある程度はそれも防ぐことができるということです。そして、オープンアクセスのモデルなので、ユーザーにとってはコスト的にありがたい取り組みです。誰でも最先端に近い情報が広く手に入るということで、よいところがあるため、化学分野でもこれをやっていこうという流れになりつつあります。

査読期間については、今まで化学の分野ではそれほど問題になっていませんでした。追試がしやすく、バイオ系に比べたら実験期間もそれほど長いわけではないので、リバイズもそこまで時間がかからなかったのです。けれども、最近様変わりしてきているのではないかという実例を挙げたいと思います。図 12 は私が関わった論文で、ファーストオーサーが ERATO のファンドで雇われていたポスドクの方で、星が付いている方はグループリーダーです。ERATO は 5 年間のファンドなので、5 年後には 2 人ともどこかに職を探さなければいけません。当然ながら追い詰められながら仕事をします。3 年ぐらいのときに論文が出せそうだということで、Nature Chemistry に投稿しました。そこからレフェリーに相当シビアなリバイズ要求を送られました。リバイズ 3 ぐらいまで行ったと思うのですが、結果として、レシーブが 2014 年 9 月、アクセプトが 2016 年 5 月で、ほとんど 2 年近い期間がかかってしまい、最終的に論文が通ったのがほとんど任期ぎりぎりという、考えたくもないような状況になりました。

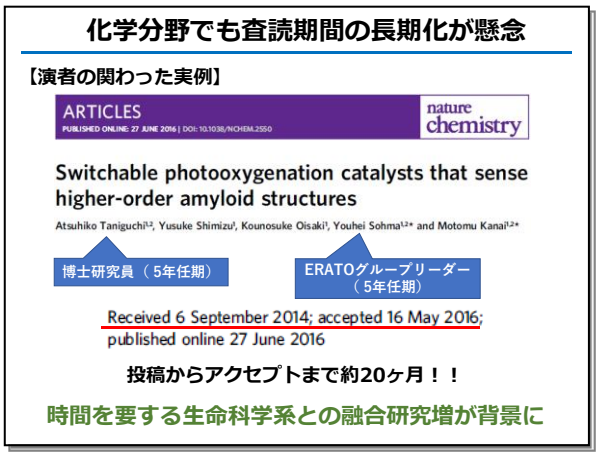
なぜこうなってしまったかというと、この仕事は純粋な化学だけではなく、細胞を使うなど、バイオロジーと融合した面もあり、時間のかかる実験がたくさんあって、リバイズも時間がかかってしまったからです。要求される実験量も多いという事情が化学の世界にもたくさん出てきたことが、査読の時間が長くなっている背景にあるのではないかと個人的には考えています。

図 13 は、合成化学のトップラボのスクリプス研究所の研究者が投稿した、複雑な化合物を斬新な考え方で化学合成したという論文です。

図 14 は論文を ChemRxiv で見たときのブラウザ画面です。論文テキストがあって、下の方にダウンロー



（図 11）



（図 12）

ドボタンがあります。更にスクロールしていくと、もう少し詳細な情報がある画面に行きます（図 15）。ここでポイントになるのが、左上にある「Version2」というものです。プレプリントで未完成品ということなので、適宜修正が入って、全部記録されて残るようになっています。

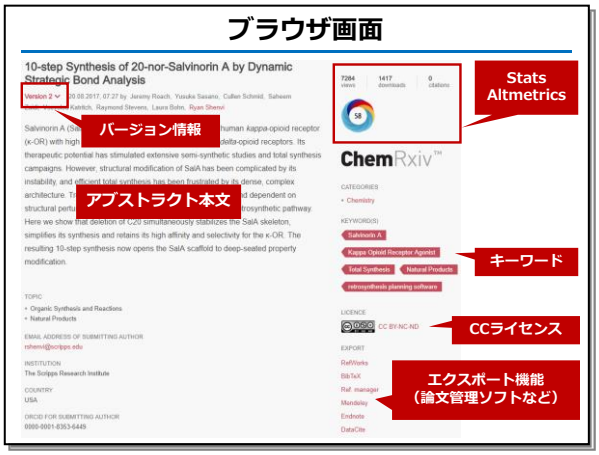
左がアブストラクト本文で、右上がどのようなソーシャル・ネットワーク・サービスで見られたかという、Altmetrics という指標で評価されるスタッツです。あとはキーワード、CC ライセンス、エクスポート機能などよくある機能が付いています。

論文の査読をスキップして、専門の研究者、現場の研究者がこういうものを見られるようになったというのは、化学の世界では革命的なことだと思っています。プロの研究者としてどのような使い方をするか考えてみると、一流ラボの論文投稿プロセスが分かるということが非常に勉強になる点です。例えば、こういう題材のものは最初にどういうジャーナルに出すのか、そこにアクセプトされなかったらどういう対策を取っていくのか、それがリビジョンなどを追っていけばある程度分かるようになるということで、私たちからしてみると非常にありがたい情報の宝庫です。

また、査読をする人の感覚も分かると思います。査読をした人が競争相手だったので理不尽に負けたり、「私の論文を引いていないではないか」と言って、無理やり引用文献を付けさせたりということも実はたまにあるのですが、そういうものもリバイズを見ていけば分かるかもしれないというので、今後は査読側も考えてコメントするようになるのではないかと思います。無茶苦茶な査読も、ひょっとしたらこういうものがあることによって減っていくことが考えられます。

（図 13）

（図 14）



（図 15）

3-3.プレプリント投稿の注意点

図 16 のプレプリントは最初、「files under embargo」という設定になっており、その間は、誰もこの論文の中身にアクセスできないようになっていました。「プレプリントなのに全部隠すの？」と私は思っていたのですが、後々オープンになったのです。

オープンになって論文のフォーマットを見てみると、「Journal of the American Chemical Society」というアメリカ化学会（ACS）が出しているフラッグシップジャーナルに投稿していたことが分かりました。著者はなぜこういうことをしなくてはいけなかったのかということですが、ジャーナルごとにプレプリントの取り扱い方が大きく違うことが原因かと思います。

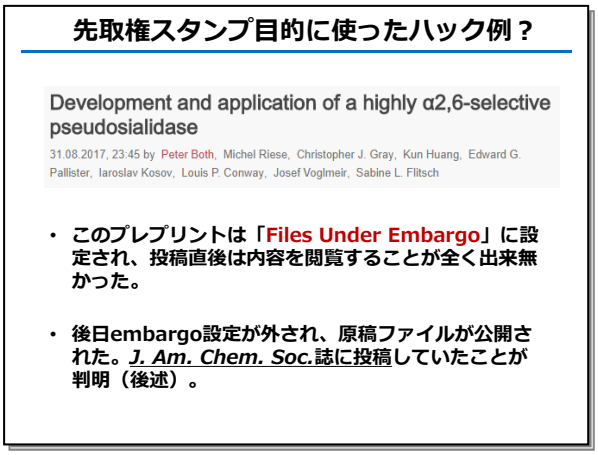
図 17 は論文投稿規定です。例えば、「Nature Chemistry」は本当に有名な化学雑誌ですが、この論文投稿規定には、「unrefereed web preprints do not compromise novelty」とあるので、プレプリントを出しても新奇性は毀損されません。

一方で、「Journal of the American Chemical Society」の論文投稿規定には、「that no portion of this or any other closely related work is under consideration for publication elsewhere in any medium」「publicly accessible preprint Web sites」とあり、プレプリントを出したら投稿済み扱いにするということです。「Journal of the American Chemical Society」に出している論文はアクセスできるような状態にしてあってはいけないという投稿規定があるので、先ほどの著者は非公開にしていたのだろうということです。

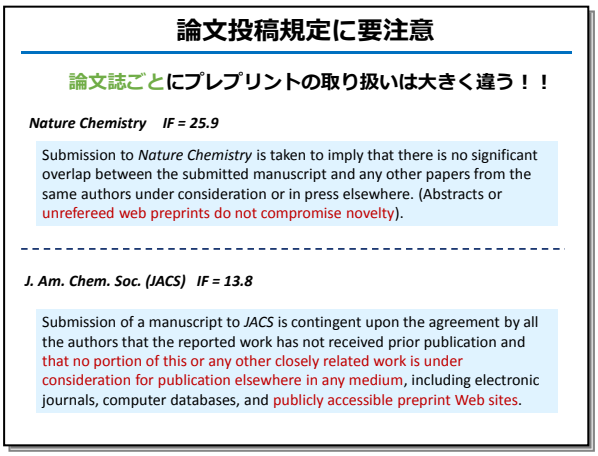
論文投稿規定は出版社単位ではなく、ジャーナル単位で違います（図 18）。ACS は「Journal of the American Chemical Society」以外にもジャーナルを出していて、その「ACS Catalysis」は触媒分野なのですが、これに関しては、「プレプリントサーバーに投稿しても大丈夫です、その代わりプレプリントサーバーに出したということをカバーレターなどでエディターに知らせてください」という一文が付いています。

ですから、同じ出版社の管轄ジャーナルであっても、扱いが全然違うということがあり得るので、それを頭に入れた上で投稿規定をきちんと読んで、投稿戦略を練らなければといけないというのが研究者視点からの話です。

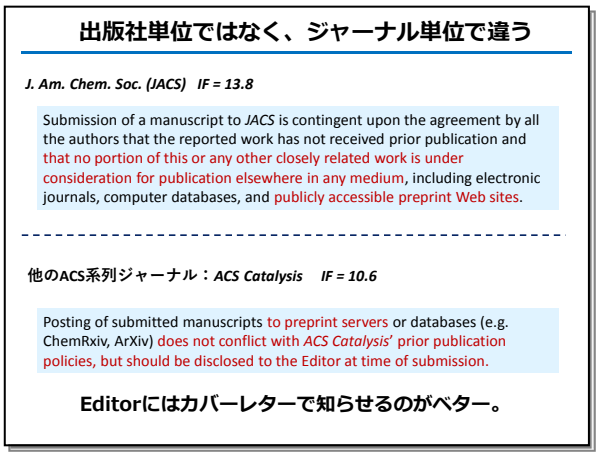
考えなしにプレプリントを投稿してしまうと痛い目を見かねないという例を挙げます（図 19）。「Nature



（図 16）

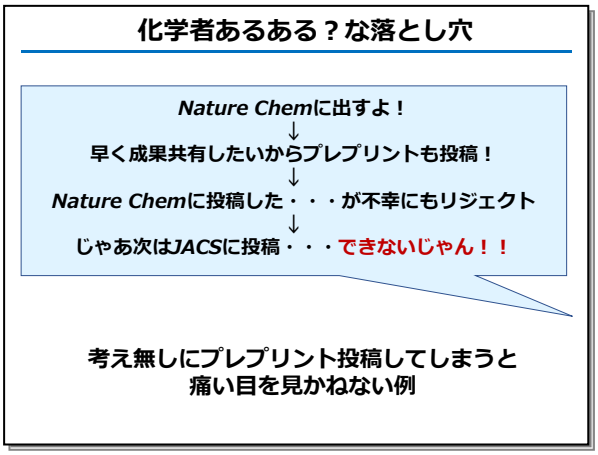


（図 17）

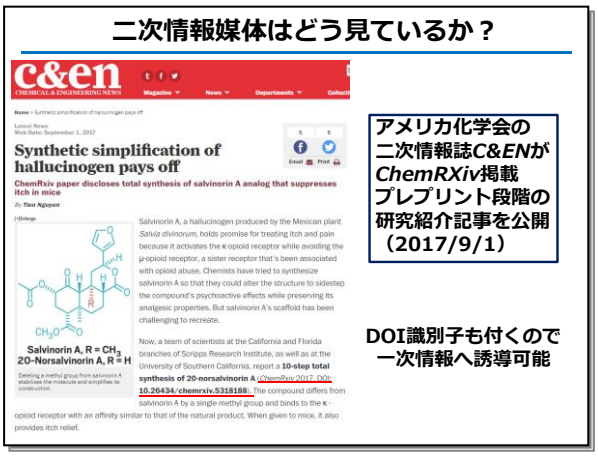


（図 18）

Chemistry」はプレプリントを出しても大丈夫なので、「Nature Chemistry」に投稿して、早く成果共有もしたいからプレプリントも投稿しました。でも、「Nature Chemistry」には不幸にもリジェクトされました。では別の雑誌にということで「Journal of the American Chemical Society」に投稿しようとしたら、プレプリントを出してしまっているので投稿できないということが往々にして起こるのではないでしょうか。何も知らずにプレプリントを素晴らしいと言って投稿していると、結局損を見るのは自分だということがあり得るので、きちんと考えて使った方がいいという段階にあるのが化学分野の現状だと思います。



（図 19）



（図 20）

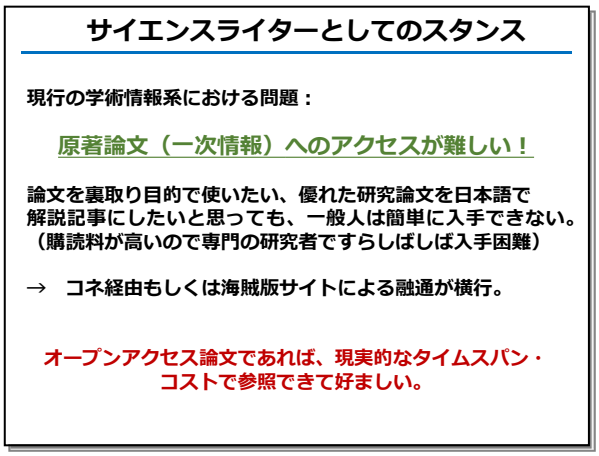
3-4.二次情報媒体はプレプリントをどう見ているか

私は Chem-Station をやってきて、サイエンスライター、サイエンスコミュニケーションという視点からものを見るようになったということもあり、実際の論文を分かりやすく紹介する二次情報媒体が、このプレプリントをどう見ているかについてお話しします。

「C&EN」という ACS の二次情報誌は、ChemRxiv に掲載された、論文になっていないプレプリント段階の研究を、2017 年 9 月 1 日に、素晴らしい仕事だと紹介して宣伝しています（図 20）。DOI 識別子も付くので、皆さんここに飛んで見てくださいねということも言えるようになっている仕組みです。

サイエンスライターの感覚からしてみると、プレプリント段階でも宣伝 OK とするスタンスは合理的だと思います。というのも、サイエンスライターは、研究自体の正確性を追究するのが仕事ではなく、研究というものを分かりやすく面白く伝えて、みんなの興味を引くことが仕事だからです。長いとみんな読んでくれないので、興味を引くには、ある程度簡単に文章をつくってお届けしなければいけない。でも、正確性をひたすら突っ込まれると困るということで、みんなそこは頭を使って苦労してやっています。

そういう事情があるので、プレプリントなのだけれど、プレプリント以降のものと比較してどこまで違うのかというと、プレプリント段階でもサイエンスライター的には十分使っていい情報なのではないかという



（図 21）

感覚になるのも無理はないと思います。本当にそれ以上のレベルの厳密さが必要であれば、専門家に監修してもらうというのが合理的だと思います。

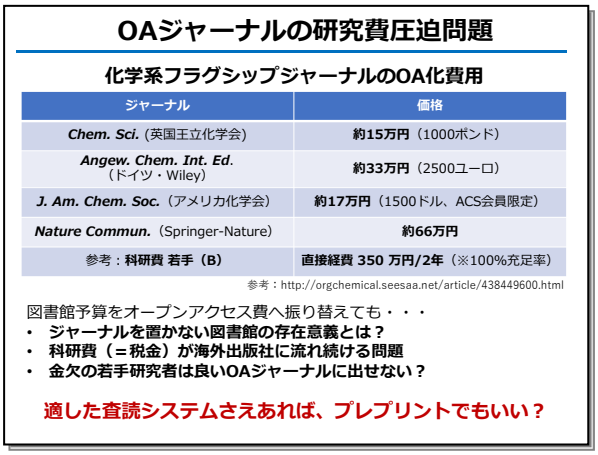
サイエンスライター的に少しきついなと思うのが、原著論文（一次情報）へのアクセスができないということです（図 21）。これは本当に大変で、サイエンスライターも正確性に気を使うのですが、正確な情報を得たいと思っても論文に手が届かないということで、やはりオープンアクセスの論文があるといいというのがみんな思っていることではないかと思います。

でも、オープンアクセスジャーナルに出すのは結構大変だというのを知ってほしいので、表をつくりました（図 22）。これは化学系の代表的なジャーナルのゴールドオープンアクセス化費用です。「Nature Communications」がとてもお高くて約 66 万円です。参考額として、私たちが取れるような若手の科研費の直接経費は 2 年間で 350 万円です。これに 1 本出すと 1 年間の研究費の 3 分の 1 が吹っ飛ぶということで、事実上これはもう選択肢に入ってきません。でも、これはよくプレスリリースなどになり、良いジャーナルであったりしますから、若手は困ったなということになってしまうと思います。

この辺のお金の配分も考える余地があると思うのですが、事実上プレプリントで上げて、きちんとした査読システムを通るのであれば、プレプリントでいいのではないかと私たちは思わなくもありません。そろそろ良い査読システムができないかということで、いろいろなところで取り組みをしています。

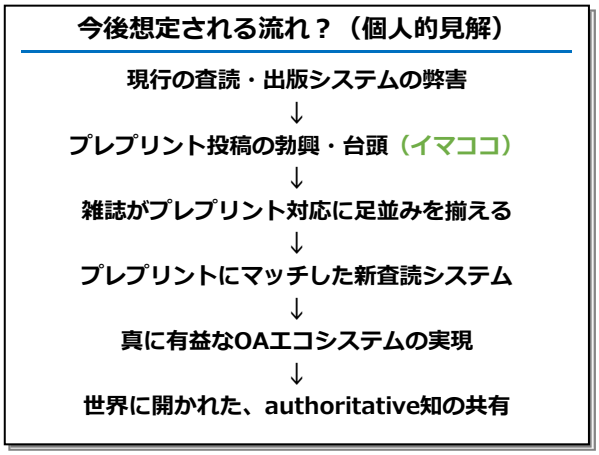
最近、「Synlett」という有機合成化学のジャーナルが、インテリジェント・クラウド査読というものを始めました（図 23）。100 人ぐらいの査読者にクラウドで議論させて結論を出すというものです。これで結構早く片付くのではないかというアイデアです。レビュアーをきちんと選ぶことが重要らしいです。エディターはきちんと誰が査読しているか分かっているので、査読者が変なことを言わないようにしているということが重要なようです。

最後の結論です（図 24）。今のところプレプリント投稿の勃興・台頭まで来ているので、ジャーナルごとに対応が違うと困るので、ジャーナルがプレプリント対応の足並みをそろえ、そこからプレプリントに合った査読システムができてくれば、いい感じでわれわれの専門知が世の中に広がっていく世界が出来上がるのではないかと考えています。



（図 22）

（図 23）



（図 24）

●**フロア 1**　プレプリントを普及させるという観点でのお考えをお伺いしたいです。bioRxiv にしても ChemRxiv にしても、それなりにサイエンスの分野では受け入れられていると思うのですが、国内の日本語雑誌の話ですが、私自身、日本語の解説記述というかオピニオンのようなものを、編集者とやりとりして、プレプリントとして公開したものも投稿するというので、figshare を使って出したことがあります。その際に編集者の方がものすごく懸念されていたことがあります。プレプリントなので「ここに投稿します」と書くのはいいだろう、でも、国内雑誌で日本語なので、いわゆる研究家の方といいますか、別に研究職に就いていない方も自由に学会員であれば投稿できるというもので、サイエンティフィックには疑問があるものも非常に投稿が多いという問題があります。プレプリントで、査読済み学術論文とは異なることを世間の人は知らないわけです。「この雑誌、この学会のものに投稿中です」と言って出して、それが広く流通してしまう。そして、その投稿中というものを、日本のサイエンスライターが一次情報にアクセスしにくいというのは割と話題になっていて、非常に懸念されていたのです。

こういう懸念というのは、理屈上は bioRxiv にしても ChemRxiv にしてもある程度、はねるとはいえあり得ると思うのです。今後、プレプリントが広がる上でそういう問題が発生し得ると思うのですが、そのあたりについてお考えがありましたら聞かせてください。

●**生長**　クオリティの低い論文が流通して、それを受け止める方が間違って受け止めてしまうことによる弊害というお話だと思うのですが、当然それはあり得るのではないかと思います。研究者のような専門知識を持っている方であれば、恐らくそこを補正しながら使うことができる、別にプレプリントでも普通に査読を経て出ている論文でも、はなからは信用しないという目線で評価できると思うのですが、では、そうでない方がどうやって使うかということかと思います。

そこは、それぞれの方に、この情報源はこれぐらいの信頼度を持っているということを理解していただいた上で使っていくしかないものなのではないかと個人的には思っています。ですので、例えばサイエンスライターの方でも、一次文献にアクセスできるのであれば、そちらを情報源として重く捉えて文章に書き加えたり、プレプリントであれば、この段階の情報なのだけれどと言外ににおわせるような表現を使って書いたりするのが誠実な対応なのではないかと思います。

●**フロア 1**　私は化学の分野ではないですが、化学の分野でこの ChemRxiv は始まったばかりとはいえ、プレプリント業界で類似の課題は少なくとも話題になっていないような気がするのです。実際の事例はありますか。

●**生長**　そもそもプレプリントを専門家以外の人が使うというケースがそこまで多くないのかもしれません。また、日本ではまだサイエンスライターがそこまで市民権を得ている職業ではないというのが一つ大きな問題かと思います。これからサイエンスライターが稼いで食べていけるような職業になるのであれば、恐らくは長期的に見て絶対出てくる問題かと思いますが、現段階でそこまで問題になっていないのではというのが私の見方です。

●**フロア 2**　NISTEP の林（SPARC Japan 運営委員）です。今の議論で思ったのですが、プレプリントにライターが先に価値を付けて記事を書くというのは、査読に加わる、あるいは査読に影響を与え得るのではないかという気がします。ピアレビューが終わる前に「この論文、すごい」と言ってしまったら、それは一つのクオリティアセスメントですよね。そうすると今後ライターがピアレビュー的なものにコミットするような

社会が、日本はまだ構造的な問題でなっていませんが、アメリカなどでは進むような気がするのですが、いかがですか。

●**生長**　そういうことはあり得る話かと私個人も思っています。Gordon さんが講演で言っておられたように、サイエンスに取り組む人がフラットになるという思想につながっていくのではないかと思います。

もちろんサイエンスライターだってある程度の専門性を持ってやってきていて、全部に全部はコメントできないかもしれませんが、例えばここに関してはすごく鋭いコメントをするという人がたくさんいていいはずなのです。サイエンスライターの出自のようなものをきっちり管理して反映させるようなシステムがあれば、そういうことはあり得る話かと思います。

●**フロア 2**　却下されるようなプレプリントを取り上げてしまうと食べていけなくなるので、サイエンスライターのフィルタリングの機能というか、ブランディングにもなるのではないでしょうか。

●**生長**　そうだと思います。また、論文の価値を評価できるサイエンスライターなのだと言えるようになる道具にはなるかと思います。

●**フロア 3**　NII の北本と申します。今の質問と関係するのですが、私も昨年、プレプリントサーバーの話を SPARC Japan でしました。やはりプレプリントは玉石混交なので、目利きが必要になると思います。逆にその目利きのブログを購読するようなモデルができるのではないかと思っています。その目利きがサイエンスライターになるのかもしれないのですが、目利きが玉を選ぶというところで何か新しいビジネスモデルができたらと思うのですが、そういう構想はありますか。

●**生長**　先ほどの「C&EN」などは本当にそんな感じではないかと思います。プレプリントの段階で、大々的に、素晴らしい研究室の誰それが仕事をして出していると宣伝しています。厳密に言うとこれは査読を経ていないので論文としてカウントされないかもしれないのですが、ACS としてはそれが合理的だと判断したからそのようにやっているところがあるのだと思います。

素晴らしい若手研究者を前に出すことによって、ACS 全体に益があるから「C&EN」がそれを担っていて、「C&EN」も Ph.D.を持っていて論文の目利きができる編集者がたくさんいる、そういう事情があるからできる話なのではないかと思います。そういうビジネスモデルを現実的なものにしていこうとすると、おっしゃるとおり目利きができる人材の層をもっと厚くしないとかなり厳しいと思っています。今後どうなっていくかと言われると、特に日本ではどうかなというのが個人的な懸念事項ではあります。

●**坊農**　私から一つケムステの方で聞きたいことがあります。私からしたら、大きなコミュニティに育っていると思うのですが、ここまで大きくなった要因とは何なのでしょうか。一つは、日本化学会と楽しそうなことをされたりしていることかと思います。われわれの生物系学会では、「特定の団体と結び付いて」みたいなことを言われて、なかなか実現しないのですが、そちらの方でいい知恵があれば聞かせてください。

●**生長**　当時の日本化学会会長だった玉尾先生がそういうことにすごく頭の柔らかい方だったというのが、一番大きな理由だったのではないでしょうか。もちろんそれまでにある程度ユーザーやリーダーの数を稼いでいないと、また、絶大な学生からの信頼を勝ち得ておかないとそうはならなかったとは思います。

それは、私たちは兼業で十何年とやってきて、進みが遅いながらもブランディングを重ねてきたというところがあります。もう一つ、そのブランディングに信頼性を与えているものがあるとすると、代表にしろ私にしろ、研究者が本業で、教育にもコミットしないと

いけない人材である大学教員という肩書きを持った人がやっているということだと思います。適当にやっているわけではないというのを言外で感じていただけているところもあるのではないでしょうか。

●**フロア 2**　化学会にいた人間として言わせてください。なぜかというと、代表、副代表とも素晴らしい研究をしているからなのです。2013 年に代表の山口先生は日本化学会進歩賞も受賞しています。学会の中でも研究者として非常に優れたお二人だから信頼性があって、広がっているのです。多分、本人は言いづらいと思うので（笑）。

●**坊農**　補足をどうもありがとうございます。そこでちょうど時間になりましたので、これで終わりにしたいと思います。どうもありがとうございました。

第 2 回 SPARC Japan セミナー2017

「プレプリントとオープンアクセス」

生命科学分野におけるプレプリントの位置付けや経験について，統合 TV について

小野 浩雅

（情報・システム研究機構 ライフサイエンス統合データベースセンター）

講演要旨

生命科学分野では、2013 年にはコールド・スプリング・ハーバー研究所による「bioRxiv」が立ち上がったことを契機に、ここ数年でプレプリントの活用が進んでいる。査読前の論文をあらかじめオープンに共有することで、特に、日々刻々と変化する研究情勢にキャッチアップする必要のある研究者にとって必要不可欠な基盤となりつつあり、周囲の研究者の間では、プレプリントの話題になることが増えてきている。演者自身はプレプリント投稿の体験はまだ無いが、論文投稿の際にあらかじめオープンレポジトリに全てのデータを登録・公開する経験をし、今後このようなプロセスが標準となるのではないかと考えている。

小野 浩雅

日本大学大学院生物資源科学研究科博士後期課程単位取得退学。在籍中は哺乳類細胞における脱分化機構の網羅的解析をテーマに研究。2010年より情報・システム研究機構(ROIS)ライフサイエンス統合データベースセンター(DBCLS) に勤務。生命科学分野の有用なDBの使い方を動画で紹介する「統合TV」の編集や遺伝子発現等の大規模データ解析および可視化、活用支援を行う。2012年6月より現職（特任助教）。博士（生物資源科学）。

先ほどの生長さんが化学分野だったので、私は生命科学分野のプレプリントの現状等についてご紹介したいと思います。

1.自己紹介

今はデータベースセンターでコンピュータを相棒にして仕事をしていますが、学部・大学院時代は、細胞などを相手にして顕微鏡を見ながら実験をするというスタイルで研究を進めてきました（図 1)。それを、生物業界だけかもしれませんが、wet 研究といいます。一方、コンピュータをメインに使って生物学の謎を解いていくことを dry 研究といって、最近は wet と dry の融合という話が頻繁に出てきます。

そういう意味では、私はもともと wet で仕事をしてきたというバックグラウンドがあります。皆さんもたくさんお持ちだと思いますが、脂を貯める脂肪細胞を

（図 1）

他の細胞に変えるという研究をしていました。一番分かりやすい例は山中伸弥先生の iPS 細胞です。あれも結局は分化転換といって、全能性のある他の細胞に変えるという研究でノーベル賞を取られたわけです。そこの一分野と捉えていただければと思います。

この分化転換、脱分化という現象は不思議な現象なのですが、十数年前にそこを追い掛けるにはコンピュータの力を借りないといけないという状況に直面しました。縁があって、今日のモデレーターである坊農さんのところに、コンピュータで実験する手法を学びに行ったというのが私の最初の出会いです。博士後期課程のときに、コンピュータを使ってこういう生命現象を解いていくということに取り組みました。

そして、データベースを使って生物学を解いていくことを推進するライフサイエンス統合データベースセンターが、ちょうど私が博士の学位を取るぐらいのときに出来上がり、面白そうだなと思いました。周りには優秀な wet の研究者がたくさんいたのですが、こういう研究をする研究者はあまりいないのではないかと思って飛び込んだのがきっかけです。

私は学生のころからアルバイトとしてライフサイエンス統合データベースセンターに参画していたのですが、当時、ライフサイエンス研究の有用なデータベースやウェブツールが出はじめてきていて、私はそれを動画で紹介する試み「統合 TV」を始めました。自分で動画を編集して、ストーリーをつくって公開していきました。

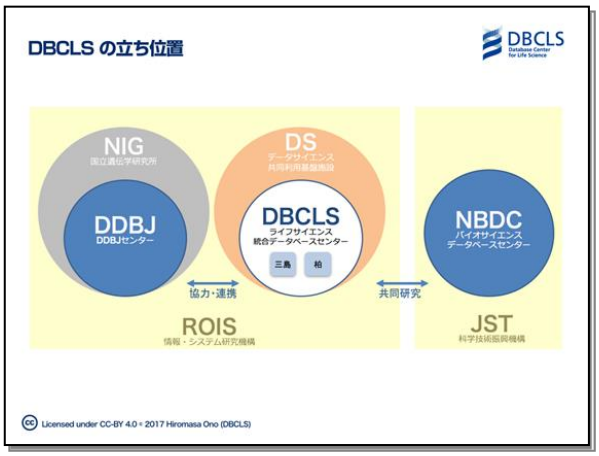
今年の 7 月か 8 月で開始から丸 10 年たち、動画が今 1,300 本に届こうというぐらい貯まっています。私がアルバイトで始めた当時は、1 日行って 1 本動画をつくるという形でしたが、今はいろいろなお手伝いをしてくださる方のご協力を頂いて、私自身はつくらず、編集者として、きちんとした内容になっているか、分かりやすくなっているかをチェックしたり、ウェブサイト自体を運営したりということをしています。最近言いはじめたのですが、これはまさに生命科学分野の「オープンエデュケーション」ツールではないかと自負しています。

もう一つ、生命科学の研究者が実際にどういうデータベースがあると日々の研究活動をより進められるようになるかということを考え、便利なデータベースをつくっていくことも、われわれのセンターの役目の一つです。われわれは実際に実験をしているわけではないのでデータを出すことはできないのですが、既に世界中に便利なオープンデータはたくさんあるので、それを再利用して、例えば、正常組織や細胞株の遺伝子発現データを簡単に検索できるデータベースなどをつくっています。クリックしていくだけで、例えばある遺伝子が心臓でどれぐらい出ているというのが簡単に分かります。このデータベースを RefEx と呼んでいます。

また、生命科学分野のデータベースをどうやって活用していくかというのを、われわれは基本的にはウェブで発信していくのですが、やはりフェース・ツー・フェースで会ってハンズオンにしないと、「これは便利だから使ってみよう」という話にはなかなかならないところがあるので、大学や研究所に出張して、データベースの講習会をよくやっています。そこで講師としてデータベース活用に関する教育もしています。

2.DBCLS の立ち位置

ライフサイエンス統合データベースセンターは、長いので DBCLS と呼ばれています（図 2）。国立情報学研究所と同様、情報・システム研究機構（ROIS）の



（図 2）

中にあります。似たような研究施設として、国立遺伝学研究所が静岡県三島市にあり、そこと兄弟研究所という形になっています。

DBCLS は三島と柏にラボがあります。DNA を研究者が解析したときは必ずどこかに登録しないといけないのですが、その日本における元締めが DDBJ という DNA のデータバンクで、遺伝学研究所の中にあります。DDBJ は毎日ものすごい量のデータが集まってくる施設です。われわれはデータベースセンターなので、これらの「ビッグデータ」を活用するために、三島にもラボを構えて DDBJ と密に連携しています。

もう一つ似たような名前のバイオサイエンスデータベースセンター（NBDC）というものがあります。NBDC は科学技術振興機構（JST）の傘下にあって、DBCLS はここと共同研究をしています。

謎の組織 DBCLS をどう紹介するのが適切かと考えたのですが、このセミナーは図書館系の方が今回多いと聞いたので、皆さん FAIR はご存じですよね。あの FAIR が実は、DBCLS の活動の中で出てきたと言ったら、どうですか。DBCLS は筋がいいなと思われた方がおられれば幸いです。見つけやすくて（Findable）、アクセスしやすくて（Accessible）、相互運用できて（Interoperable）、再利用できる（Reusable）、の頭文字を取った標語です。これについて書かれた論文が Scientific Data に出ています。

（https://www.nature.com/articles/sdata201618）

このファーストオーサーである Mark Wilkinson さんが DBCLS の片山俊明さんと仲が良く、二人の長年の構想で実現したのが、10 年前から行っている BioHackathon（バイオハッカソン）というイベントです。今年で 10 回目を迎えました。

Hackathon という言葉は最近 IT 系などでよく出てきます。Hack と marathon の造語で、1 週間ぐらい合宿で hacking し続けるのです。ハッキングとは悪い方のハッキングではなくて、コンピュータで仕事をすることを hack といいます。昼は議論しながら仕事をして、夜はお酒も飲みながら議論する、それを 1 週間ぶっ続けでやるイベントです。ここで FAIR という概念が大いに議論されたということがこの論文にも書いてあって、Acknowledgements に、NBCD/DBCLS BioHackathon 2015 のおかげでできたというようなことが書いてあります。そういう組織、集団だと思っていただければよいと思います。

3.統合 TV

3-1.概要

次に、統合 TV について紹介します。統合 TV は、生命科学分野の使えるデータベースやツールの使い方を動画で手取り足取り解説するウェブサイトで、図 3 のようなインターフェースになっています。生命科学研究を進める上で最低限知っておかなければいけないバイオ系のデータベースを網羅しています。

生命科学分野には、バイオインフォマティクスといって、コンピュータを使って生命科学現象を解き明かしていく分野があります。この分野は、先ほど wet と dry という言葉があると紹介したと思いますが、やりたくない人はやりたくない分野で、黒い画面に緑色の文字が流れてくると思考停止してしまう人がおられるのですが、そういう人に向けて、「実は怖くないのです、簡単にコピペをするだけで緑の画面を克服できるのです」ということも紹介しています。全ての動画に DOI を付けており、引用ができるようになっています。

この動画は YouTube に全て同じものが上がってお

（図 3）

り、おなじみのインターフェースで閲覧することができるようになっています。おかげさまで再生数も右肩上がりで、90 万弱まで来ています。

3-2.講習会の動画

動画の種類は、単純にスクリーンショットを撮って動画の使い方を紹介するというのが最初だったのですが、われわれはいろいろなところで講習会をしているので、その講習会の様子を動画で撮って、後から振り返れるようにアーカイブしています（図 4）。講習会では「今日は何をやります」という形で、説明書や手順書を用意されると思うのですが、それを実際にウェブサイトに載せておいて、動画を見ながら、手順書を使って、自分の手元のコンピュータで再現できる、再学習できるという用意をしています。このページには DOI が付いていて、再利用できます。

この分野には最新の解析手法があって、今、ヒトゲノムは非常に安い値段で解析できるようになっています。1,000 ドルゲノムといわれているぐらいで、昔は 1 個解析するのに何百億円とかかっていたのに、今はそれぐらいの値段で、しかも短時間で解析ができるようになりました。その一つのキーワードが次世代シーケンサー（NGS：next generation sequencer）です。これが開発されたのが大きなブレークスルーだったのですが、この解析手法が非常に難しいのです。

生物学はコンピュータを使わないと解析ができないレベルになってきています。その最近の解析手法の講習会の動画（図 5）は、コンピュータリテラシーとサーバ設計の話から始まっており、そういうレベルのデータを扱う必要が生命科学者の研究者の中にも出てきたということを意味しています。後の方になると生物学的な話になってきます。

この講習会は 3 年前に第 1 回があって、2014 年分は延べ 43 時間を超えています。43 時間も YouTube の動画をご覧になったことがありますか。1 本は 2 時間半や 3 時間で、これでも黙っている時間などをカットとしているのです。この講習会は、1 週間通しの集中講義形式で東大農学部で行いました。

NGS で職を得ることはできると思いますが、プロとしてやっていこうということであれば、番組を最初から最後まで見てやってみて、やっていけそうかどうかを試してみるといいと紹介しています。毎年、時間数が増えているのですが、今年もあって、絶賛編集中です。

3-3.生命科学分野の静止画素材

動画だけではなく、静止画の素材も統合 TV で提供しています。もともとは Togo Picture Gallery という名前の別のサービスでした。しかし、ウェブサイトが別だとなかなか普通の人はたどり着けないというご意見を頂き、それだったら有名な方の統合 TV にくっつけようということで、1 年ぐらい前に合体しました。

これもわれわれのセンターでコンテンツをどんどんつくっていくために、RA さんを呼んで手伝ってもら

（図 4）

（図 5）

いました。ある学生が、動画をつくるのはあまり上手ではなかったのですが、絵を描くのが非常に得意で、「絵でもいいから描いて」という軽い気持ちで始まったのですが、今や 500 本ぐらい素材がたまっています。これはクレジットを明記するだけで、どのように使っても、商売に使っても構いません。誰でも自由に、閲覧するだけではなく、使うことができる、ライフサイエンス分野の画像やイラストです。種類も生命科学分野に何かしら引っ掛かっていればよいというポリシーでやっているので、各種取りそろえていて、研究発表のスライドや資料などに使っていただけます。よく広告の挿絵に使いたいというお問い合わせも頂いています。

最近は細胞やマウスなどの 3D アニメーションも取り込んでいます。3D プリンターをお持ちの方はなかなかおられないと思うのですが、最近、街で 3D プリンターを貸し出しているところもありますので、使っていただくと楽しいと思います。

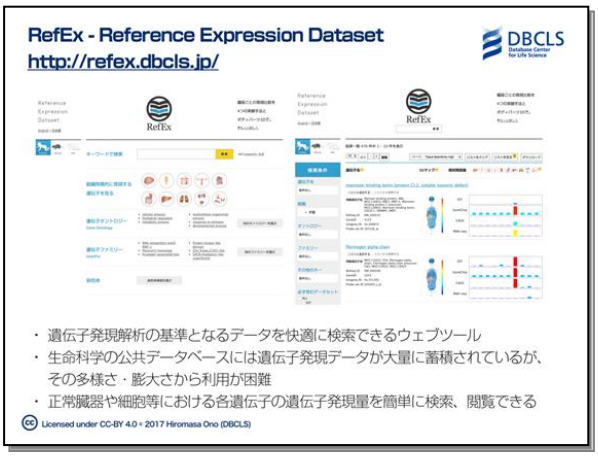
3-4.RefEx（Reference Expression Dataset）

冒頭で紹介した RefEx は、正常組織や細胞の遺伝子の発現量を簡単に検索できるサービスです（図 6）。遺伝子発現解析の基準となるデータを快適に検索しようというものです。キーワード検索もできますし、「組織特異的に発現する遺伝子を見る」という欄から臓器の絵をクリックして検索することもできます。

2017 年 8 月の終わりぐらいに、この RefEx に関する論文が出ました。奇しくも FAIR の論文が出たのと同じ Scientific Data というデータジャーナルにです（図 7）。これはまさにオープンアクセスです。私自身は実はプレプリントを投稿したことはまだないのですが、データジャーナルに投稿した経験を話したいと思います。

この Scientific Data 自体も生命科学分野で珍しく、普通のジャーナルとは違い、よりデータにフォーカスを当てて紹介するという、チャレンジングな試みです。もともと、研究者が出した、既存のデータベースに載っている一次データセットを他の人が再利用しようとするとき、適切に再利用するための情報が少ないという声がありました。それを一生懸命書いても一次データを出す人にとっては何の業績にもならないからです。これでは Win-Win ではないということで、データジャーナルが出てきたのだと考えています。まさに「一次データセットについて測定の対象、方法、品質を記述する」ために始まったジャーナルだったのです。

しかし、われわれは自分自身でデータを出しているわけではなく、それを再解析して使えるウェブツールをつくっているので、この規定とはそぐわないところがあったのですが、ちょうど運良くタイミングが合い、Scientific Data が再利用を決定的に促進するシステムや技術についての独自のレポートとしてアーティクルタイプの論文を新しく受け付けることになりました。ただ、アーティクルで出すときも、Data Descriptor と同じように、使ったデータについてしっかり記述しなけ



（図 6）

（図 7）

ればなりません。

再解析したデータを全て DOI 付きで、全ての人が参照できるところに出さなければいけないということで、われわれが使った全てのデータを figshare に公開し、43 個ぐらいのデータセット全てに DOI を付けました（図 8）。

また、どうやって再解析をしたかというプログラムも全てつまびらかにしなければいけないということで、プログラマーが使うリポジトリのサイト GitHub に、公開データの再解析に用いたプログラムや何をどう処理したかという文書をきちんと整理して置きました（図 9）。

4.生命科学分野のプレプリントサーバ

次はプレプリントサーバについてです。生命科学分野のプレプリントサーバは、bioRxiv が非常に有名で皆さんご存じと思います。一番初めのプレプリントサーバは arXiv.org でしたが、arXiv.org に生命科学分野のものが全て入っているかというとそうではなくて、quantitative biology といって、一部分の生物物理寄りのものだけ入っていました。生命科学全般に対応したようなものはあまりなかったのですが、今は図 10 のようなものがいろいろ出てきているという状況です。

プレプリントサーバを検索できるサービスも出てきています。図 11 の PrePubMed は、まさに PubMed に載る前のプレプリントを、各種プレプリントサーバ横断的に検索できるサービスです。PrePubMed に面白い統計が載っていたので引っ張ってきました（図 12）。今年 9 月の最新のプレプリントの統計情報で、一月で約 1,500 本のプレプリントが追加され、オーサーも約 750 人増えているということです。

このグラフで非常に面白いと思ったのは、緑色が

（図 8）

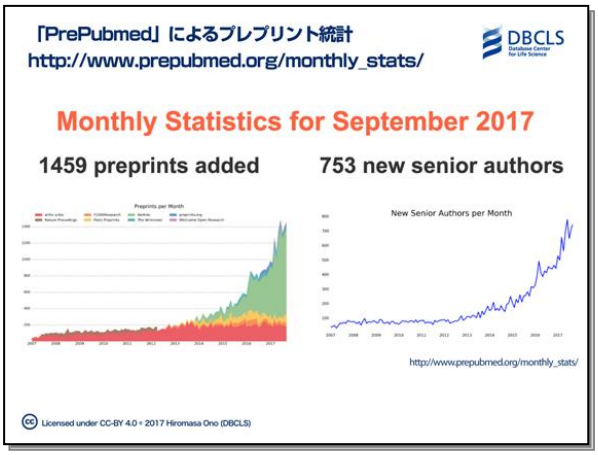
（図 10）

（図 9）

（図 11）

bioRxiv で、面積が急激に増えていることです。bioRxiv が出てきたのを契機にプレプリントがはやってきたということが分かると思います。オーサーも、ちょうど 2014 年ぐらいに増えてきたということが見受けられます。

bioRxiv は 2013 年 11 月にアメリカのコールド・スプリング・ハーバー研究所が開始しました。投稿自体は自由なのですが、論文の体を成していない、内容がふさわしくないというものは選別されているようです。これはどこのプレプリントでもそうだと思いますが、査読後に出版される前の研究成果の迅速な流通とフィードバックの活性化を目的としています。bioRxiv は現在までに 16,000 報のプレプリントが投稿されています。最近のアップデートで、有力な学術誌にそのまま転送して投稿できる機能も付いたそうです。e メールや RSS、Twitter でのアラートシステムが充実しているというのが特徴の一つです。



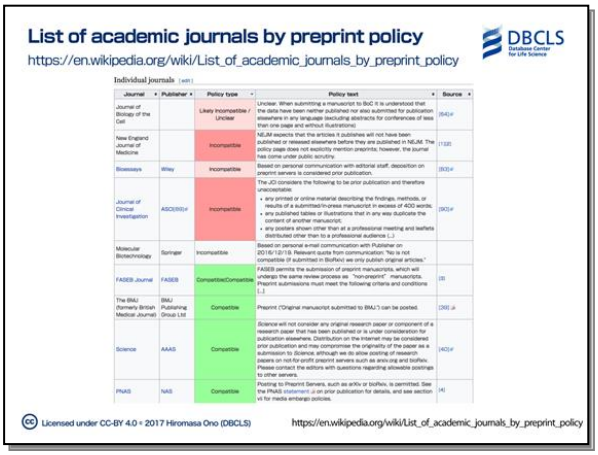
（図 12）

図 13 が生命科学分野のプレプリントのポリシーです。Wikipedia にこのようなエントリーがあって、随時更新されています。どのぐらいのジャーナルでプレプリントを認めるのか調べてみたら、プレプリントを認めないところが意外と非常に少なかったです。これらは独立系のジャーナルで、五つぐらいはまだ認めていないようですが、それ以外は compatible な状況のようです。

DBCLS の同僚の Tazro Ohta（大田達郎）さんは、図 14 のお花見メタゲノムについての論文を bioRxiv に投稿しています。お花見メタゲノムとは、お花見をして、桜の細菌叢をみんなで集めてそのゲノムを読むという面白いプロジェクトです。bioRxiv に載せているだけではなく、ジャーナルに投稿中で今リバイズをしている状況です。

bioRxiv への評価を Twitter などで見てみると、非常に好意的な方が多いです。好意的な意見を集めてきている面があるので話半分に聞いてほしいのですが、ある人は「bioRxiv にプレプリントを置いたら世界中から反響が来て、論文はまだアクセプトされていないけれど、別の国際共同研究が始まった」「有名どころが自分たちの材料を使って実験してくれていて楽しい」ということを言っています。

もう一つは、実感できる内容だと思うのですが、「読みたい雑誌の記事が有料だったから読めなかったけれど、タイトルで検索したら bioRxiv に PDF があっ



（図 13）

（図 14）

ていて読めた、ワーイ」というコメントです。このように bioRxiv を使っている方もいるという状況なのだと思います。

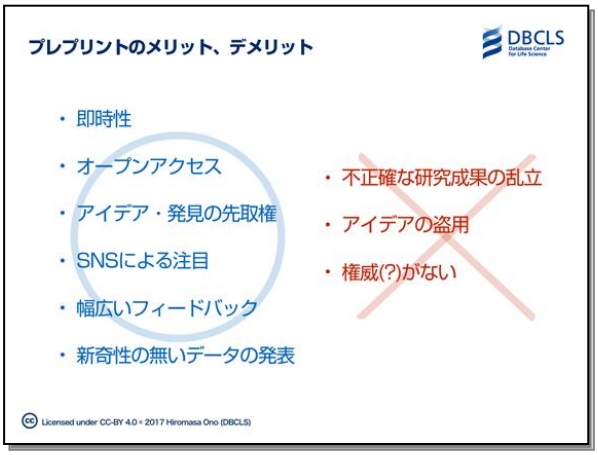
5.プレプリントのメリットとデメリット

最後に、生命科学分野に特化した話ではないと思いますが、プレプリントのメリットとデメリットを自分なりにまとめてみました（図 15）。

メリットは、まず、即時性があること、オープンアクセスであること、アイデア・発見の先取権が取れることが挙げられます。また、今はみんながスマートフォンを持ち歩いている時代で、気軽に SNS で情報が行ったり来たりしているので、注目されるというメリットが一つあると思います。さらに、それによって幅広いフィードバックが得られます。

もう一つは、新奇性のないデータを発表できるということだと思います。生命科学に限らず、オリジナリティのない論文はジャーナルベースで世にあまり出ません。再現性を確認しただけの実験や、やってみたのだけれどうまくいかなかった実験は、既存の媒体に載りづらい性格のものですが、既存のジャーナルのアーティクルの形式にのっとった形で公表できる媒体はメリットだと思います。

デメリットは、不正確な研究成果がたくさん出てきてしまう懸念があることです。また、コアのアイデアを盗まれて転用されてしまうことを心配する現場の研究者の声もあるように思います。そして、権威がないというか、今までの既存の評価軸に乗らないところがあるので使いたくないという意見もあるように聞いています。



（図 15）

●**フロア 1**　NICT の研究者です。生物はプレプリント文化を先進的に切り開いてこられたということで、学会内でどの程度コンセンサスがあるのか、若手は業績に対して安心感を持ってできているのか、研究者の日常の目線から見るとどんな感じかを共有いただけたらありがたいです。

●**小野**　私が見聞きしている範囲ですが、やはりまだ「プレプリント、ああ、何か聞いたことあるね」というレベルが平均的な生命科学系研究者の感じ方だと思います。今のところはアーリーアダプター、アンテナの感度が良い人が、先ほど紹介した Twitter もそうですが、たくさん発信もして、発信されたものも受け取っていて、「いいね」という雰囲気になっているのだという感じがしています。それが学会としてとか、普通のインフラとなっていくには 2〜3 年、あるいはそれ以上かかるかもしれません。

評価されるということに関しては、例えば科研費の申請書には、bioRxiv に挙がっているけれど、「(査読なし)」ときっと書くわけですよね。それを見た評価者が、「何だ、査読なしでは駄目だ」と評価するのか、「プレプリントを使っているとはなかなか筋があるな」と捉えるかどうかという、評価者個人がどう思っているかにまだ依存するところがあると思います。そこでどのようにコンセンサスを得ていくかは、生命科学に限らず、学会なり、国なりがどういう施策をしていくかということに依存するのではないかと思います。

●**フロア 2**　NBDC の職員です。小野さんがご存じかどうか分からないのですが、「Nature」の記事だったと思いますが、bioRxiv に載っている論文は、厳しめ

のライセンスを付けている人が多いということが載っていました。クリエイティブ・コモンズ（CC）を付けることもできるのですが、そもそも CC を付けずに普通の権利を付けていたり、CC でも特に厳しい改変禁止を付けていたりするものが非常に多いという話があります。

一方では、どうもその記事を読んだ印象では、bioRxiv のポリシーでは論文に関してはテキストマイニングが全てフリーで行えるということなのに、大半の著者はそのことを意識せずに投稿しているのではと思うのですが、その辺について聞いたことはありますか。

●**小野**　私もその話を聞いて、「へー」と思ったところです。やはりまだ bioRxiv についても、行き届いていないという批判的な記事も見つけました。プラットフォームとして今一番有名なのでみんなこれを使っていますが、まだ知らないでライセンスを付与しているところがあるので、その辺をきちんと意識してライセンスを付ける、付けないという意思表示がきちんとできるように、まだまだプラットフォームとしての改善の余地があるのではないかと思います。

●**フロア 3**　NII の北本と申します。データジャーナルへの投稿で解析済みデータを figshare に寄託したということについて、例えば DBCLS そのもので公開することはできないのかということと、日本のデータが figshare に流れてしまうことについて、こうしたらいいのではないかという意見をお願いします。

●**小野**　非常に重要なご意見ありがとうございます。われわれは公開リポジトリに出したのですが、Scientific Data が認める公開リポジトリであればどれでもいいという形になっています。NBDC のデータベースアーカイブというサービスも実は Scientific Data が認める公開リポジトリなのです。しかし、これは言っていいのか分からないですが、やはり figshare は楽なのです（笑）。これはインターフェース的な問題です。われわれも NBDC と共同研究をしているので、日本のところに置くべきだろうというご意見を頂くのはごもっともなのですが、やはりいざ論文を自分で出そうとすると便利な方に流れてしまったというところがあります。

ただ、「figshare に流れてしまう」という表現をされていましたが、彼らが囲ってしまうわけでもないですし、プラットフォームとして用意されているので、個人的にはそれでいいのかなという気持ちです。NBDC さんにはその辺を頑張っていただきたいなという思いです（笑）。

●**坊農**　補足します。この責任著者として言わせていただきますが、これを始めたときは、NBDC でそういう指定はなかったのです。だから、どこかに上げないといけなかったので、Dryad か、figshare に上げるかと調べたら、Dryad は有料、figshare は無料ということで、お金がないので figshare になりました。

●**フロア 4**　NISTEP の林（SPARC Japan 運営委員）です。GitHub の例を出されたように、オープンソース文化が特に dry なバイオサイエンス分野にあるので、私は何となく FAIR 原則はもっと文化として取り込まれているイメージを持ちがちなのですが、それでもプレプリント文化のようなものはまだ進まないのですか。情報系バイオサイエンティストと wet 系バイオサイエンティストの差のように単純に分けていいのか、もう少し複雑な研究者文化背景があるのか、教えていただけますか。

●**小野**　まさにご指摘のとおり、バイオインフォマティクスといわれている人たちは、こういう FAIR のようなものは意識していたか、していないかは別として、頭の中にはあった概念だと思います。

●**フロア 4**　GNU ライセンスなどはもとからそうで

すよね。

●**小野**　そうですね。もともとそうなので、むしろ何でオープンではないのか、というところです。私たちの周りでも、バイオインフォマティクスの人たちがかんでいるものが積極的に bioRxiv などにも投稿されています。当然、存在も知っているし、「こういうのに出していこうよ」「今、これが一番受けているんだよ」という感じでプッシュしているという印象があります。ただ純粋に、実験系のラボ、医学系のラボなどだと今のところ「何これ」という話になっていると思うのです。当然そういうことを教えてくれる人もいません。

●**フロア 4**　所属する組織の文化に引きずられるのですか。

●**小野**　そうだと思います。周りの付き合っている人の文化は多分にあります。おっしゃったように、IT 系というか、インフォマティクス系は全然これに対する抵抗はないのです。むしろ「何で出さないの」と言われるぐらいです。ただ、それが全部の分野に適用できて、公開前から出していいのかという話にも当然なるので、これが大正義というわけではないと思います。そこは純然たる深い溝がまだあると感じています。

第2回 SPARC Japan セミナー2017
「プレプリントとオープンアクセス」

全　体　議　論

坊農　秀雅　（情報・システム研究機構 ライフサイエンス統合データベースセンター）
引原　隆士　（京都大学図書館機構長/arXiv.org 会員コンソーシアム代表）
生長幸之助　（東京大学大学院薬学系研究科/化学ポータルサイト Chem-Station 副代表）
小野　浩雅　（情報・システム研究機構 ライフサイエンス統合データベースセンター）

●**坊農**　全体議論を始めたいと思います。質問し足りなかった方はぜひお話ししてください。多岐にわたってお話しいただいたので、個別の先生へのお話などもあると思います。

●**フロア 1**　産業衛生学会のものです。私はジャーナルの編集をしているのですが、海外の医学系ジャーナルでのプレプリントサーバについて、どれぐらい広まっているのかなど、状況をどなたかご存じですか。以前担当していた医学系のジャーナルで、中国からの投稿にクロスチェックをかけたところ全部一緒になって、結局プレプリントサーバに入っていたということがありました。そのジャーナルはプレプリントサーバへの投稿を認めないとしていたのですが、小野先生の資料を見ると BMJ は認めているし、以前に Elsevier の方から Elsevier ではほとんどが認めていると聞いたことがあります。海外では実際にはどんな感じかお分かりになりますか。

●**小野**　私も医学系ではないので残念ながら具体的な話は分からないのですが、Wikipedia のリストを見ると、incompatible なものは割と医学系、独立系の雑誌が多かったです。医学系が悪いわけではないのですが、保守的な感じがあるのかなと。先ほどバイオインフォマティクスの方が先に進んでいるという話がありましたが、そういう意味では、中身のセンシティブさに時代が追い付いてきていないところがあるのではないかという印象を持っています。

●**フロア 2**　慶應義塾大学の市古（SPARC Japan 運営委員）と申します。アメリカ国立衛生研究所（NIH）がプレプリントに対する方針を今年出したことによって、PrePubMed という動きになってきたのだと私は理解しています。今、保守的だという話があったので少し付け加えますが、確かに慶應義塾大学の中でどの程度 bioRxiv に投稿されているかを見たところ、臨床医学の積極的な投稿は目立たなくて、理工と医学系に関わる方、それから SFC で情報科学に関わる方と医学研究を一緒にやっている方の投稿が見受けられました。

●**坊農**　私自身も少し医学系をやっているのと、BMC Genomics でアソシエートエディターをしていたこともあるので、それで見ていると、プレプリントに出したものは、リファレンスとして挙げられることはよくあって、普通の論文のように扱えるという反面、基本的には査読を経ていないので信用できないという扱いになっていたと思います。

ただ、小野さんの出してくれた Wikipedia の表も動的に変わっていって、incompatible だったジャーナル誌が急に compatible になっていくという傾向があります。ざっくり探すときの参考情報としては Wikipedia はいいと思いますが、情報のソースが最新ではないの

で、実際のジャーナルの判断をそれぞれのところへ見に行って、正しいかどうかをよく見た方がいいと思います。ただ、劇的に変わっていっている現状があると思います。

医学系で全く同じものを出してくるというのは、論文の本数を稼ごうとしている人がやりがちで、そういうものは情報技術で対抗して取り締まることが大事だと、私自身やっていてよく思います。それをやり過ぎで逆に、前の論文と同じ手法を使った論文がはじかれたということが最近ありました。自分が前に出したものとほとんど同じということでプログラムのチェックではじかれてしまって、他のジャーナルに出さざるを得ないという事態になったこともあります。

プレプリントで出したものと全く同じなど、おかしいものはおかしいとチェックできるようになってきたという点では、良い時代になってきたのだと思います。

●フロア 3　国立極地研究所の図書館員です。引原先生にお伺いします。コーネル大学で、arXiv.org のエフォート管理をしながら、ライブラリの中で実際に回しているということですが、arXiv.org は今後、拡大路線で行こうとしていると受け取りました。そのときに、ライブラリで今後どのような運用体制を取っていくのでしょうか。また、その視点から、SSRN では商業ベースに移行したという話がありましたが、そちらの可能性についてはどう考えられているのでしょうか。

●引原　それは非常に微妙なところで、Gordon さんの講演で言及のあった、SSRN が移行した理由を Twitter で拝見したのですが、確かに経営的な問題というか、人のお金の問題は大きいです。arXiv.org の場合はタスクが増えていくのに合わせて、エフォートをあるときは増やしたり、あるときは減らしたりしています。今、従来のシステムから新しいシステムに移るときに増員しています。ファンディングに対して増員を要求してお金を付けてもらうという形を取っていて、それは経常的ではありません。いずれそれは減るので、その分が今度はチェックする人たちの雇用の方に回るであろうというのが、経営状態を見ていても分かります。

かなりの部分はオートマチックになっていますが、iThenticate に引っ掛かったものなどのケアは丁寧にやっているので、それが増えてきたときにどうなるかというのは確かに問題かもしれませんけれども、現状は回っていると思います。

●フロア 4　NISTEP の林（SPARC Japan 運営委員）です。先ほど引原先生にしなかった質問を投げたいと思います。リソースシェアという意味で、機関リポジトリがせっかく 744、今はもっと増えているはずですが、あるので、アーカイブの受け先を日本の場合は機関リポジトリにしてしまうというアイデアはどうなのでしょうか。つまり、会費を払うだけではなく、ストレージの置き場やナビゲーションのところのコントリビューションで組む、商業主義的なものへの対抗として考えられるアイデアとしてそういうものが思い浮かんだのですが。個人的な感触で構わないのですが、そういう方向性をまず示さなければいけないのではないかということが 1 点と、機関リポジトリをそもそも具体的に使うことに関してはいかがでしょうか。

●引原　今年の arXiv.org の Member Advisory Board（MAB）で、われわれ日本から、機関リポジトリについてディスカッションしてほしいと要求を出しました。会議の日の昼休み（13〜14 時半）に昼食を取りながらディスカッションしたところ、リポジトリ自体についての考え方がヨーロッパと日本とアメリカでは違うということで、日本の場合はリポジトリがプレプリントサーバ的なところがあるのでその可能性はかなりあるのですが、やはりなじまないという議論がありました。

それは機関がどこまで全体のデータベースのサポートにタスクを出せるかという問題があると思うのです。今はコントロールできるミラーサイトと Amazon のク

ラウドでやりたいというのが正直なところで、京都大学の基礎物理学研究所にあったサーバがなくなったのもそれが一つの理由です。今後はクラウドになっていったときに、日本でクラウドをどこかが管理するということは当然あると思います。けれども、現状のシステム開発をしている段階では、やはり全体をマネジメントしたいという方が強いようです。

●**フロア 2**　少し下世話な話なのですが、arXiv.org は経営的には問題ないのでしょうか。慶應義塾大学の例なのですが、arXiv.org に協力するかどうかというお金の分担の話が来て、誰がどう払うかとなったときに、図書館員が価値を示していくのがなかなか難しい状況にあります。私としては arXiv.org に協力することはいいことだろうとは考えているのですが。お金の面が大丈夫であれば学内でもう少しいろいろ議論できます。将来的にこれが成り立たなくなることがあるとすれば、図書館としては何をどうすることができるのかということをお話しいただければと思います。

●**引原**　今日はお金の議論は下世話になると思って外していたのです。ここ 2〜3 年は開発の経費のこともあって増えていますが、それはいずれ落ち着きます。ただ、2018〜2022 年のモデルとして、どれだけのユーザー負担を考えるかを議論した結果、利用率に応じた負担を頼むということで今まで来ています。

現状では日本に負担が来ていると思えるのですが、かなり考慮を要求しました。日本の場合はファンドを直接アメリカに送ることは、内閣府や文科省に聞いてみてもやはり難しいということがあって、国立情報学研究所（NII）などが取りまとめてコンソーシアムで払うという形態しかないだろうというのが向こうとの認識にありました。

それは日本だけではなくて、アジアの多くの国が多分そうだろうということを言っています。もし仮に日本のファンドを、S が付いている会社などに出していただいたら、これは大きな違いがあるかもしれません。

私の認識としては、経費を出すということと同時に、今まで arXiv.org の単なるユーザーであった日本が発言権を持って、いろいろな提案をしてきています。昨年から私はかなりしていて、それで存在感を示すことができています。arXiv.org を中心として、先々プレプリントサーバが統合していく動きになると思うのです。そのときに発言権をかなりキープできるのではないかと思っているので、ここは単にユーザーとしてのお金ということだけではなくて、次のものを得るための投資として考えた方がいいのではないかと考えています。

ですから余計に、例えば「ビューワーをもっとこうしてください」と発言したっていいですし、「メタデータもこういうのはどうですか」と提案してもいいわけです。今までは単なるユーザーだったからできなかったので、やはりそういう方向に転換した方がいいのではないかと私は思っています。

●**フロア 5**　先端医療振興財団のものです。私は医療系に携わっていて、各学会との連携もあるのですが、学会同士のデータシェアリングなどの課題があります。ただ、それぞれの学会でいろいろな考え方があるので難しいという状況だと思います。それで、日本で独自にそういうデータベースを作成されるお考えがあるかどうか、それをハーモナイズして海外に展開する構想はあるか、もしないとすればどういうところが難しいか、どの先生でも結構ですので教えていただけたらと思います。

●**坊農**　私たちライフサイエンス統合データベースセンターは、予算規模的に基本的にはライフサイエンス分野で多くの方に使ってもらえるデータベースの維持・管理ぐらいしかできません。ある分野のデータベースをつくることに関しては基本、関与できないのが現状です。なぜかというと、お金というよりも、まず人が足りないからです。そういうことをやったらいいのにと皆さんから伺い、確かにそのとおりだと思うのですが、分野オリエンテッドなデータベースはその分

野でやりたい方に始めてもらって頑張っていただいて、困ったことがあれば、われわれが技術的にお手伝いするというのが現状だと思います。

●**フロア 5**　それをもし商業ベースでどこか出版社にお願いするとしたら、非常に高額になって、実現が不可能になってきます。今ある海外のものを使うということに関しても、それぞれの学会の先生の考え方があって、ここは使うがここは使わないということで、結局ばらばらのデータベースに入ってしまって、それらが結び付かないことが今後問題になってくると思うのですが、そのあたりはいかがでしょうか。

●**引原**　データベースに関して、海外の出版社の狙いはまずデータを取得することで、次はクロス的に AI などを使ってマイニングしていくことなので、日本の学会がこぞって海外のデータベースにデータを置くことは非常に危険です。ジャーナルで負けたところを、最後データで維持しているのですが、それがぎりぎりの状態にあるということを学会の先生方に理解していただかないといけないのかもしれません。

それをどうしたらいいかというときに、クラウドでも、日本の NII を中心とした形でもいいですから、機関がデータサーバをパブリックにサポートするシステムを立ち上げなければいけないのは明らかだと思います。それをやらずに、学会が学会の維持のためにやるということは危険ではないかと思われます。アーカイブが商業出版に吸収されてしまったという例がありますが、それはアーカイブを維持するためという非常に矮小な議論の結果だと思います。そこは学会が、データのユーザーとしての研究者や企業などと協議して決めていかなければいけないのではないかと言わせていただきます。

●**坊農**　NBDC から、公開データベースではないものの話はないですか。紹介していただけないですか。

●**フロア 6**　NBDC で提供しているものとしては、私が担当している生命科学系データベースアーカイブというものがあります。これは基本的には全部公開していて、あまり突っ込んだ加工はできないのですが、ある程度再利用できるような形に持っていくことはできると思います。それとは別に、エンバーゴのようなものですが、将来公開することを前提として、プロジェクト期間内はプロジェクトの中だけで共有するというサービスも展開しているので、個別に相談していただければお力になれるかと思います。ただ、基本的には、データをこういうふうにしたいというのはそれぞれ研究コミュニティの方に主導していただきたいと思います。

●**フロア 7**　これはデータの保存ということであれば、また別の面があります。2017 年 11 月 6 日（月）、国会図書館と研究データ利活用協議会（RDUF）で、データサイエンスを巡る世界の動向におけるデータのアーカイブといった問題について議論するので、そこで続きの議論ができるかと思います。

（https://japanlinkcenter.org/rduf/events/index.html#s001）

活用するデータにはあまり適切ではありませんが、データを保存するだけであれば国会図書館も使えるかもしれません。要するに活用するデータはそういうのはあまり適切ではないと思うのです。データとしてとにかく保存したいというのであれば、またアーカイブという視点では別の策があるのではないかと思います。

●**フロア 4**　今日しなければいけない大事な議論は、どうやってプレプリントが各研究者コミュニティに受け入れられるかということです。なぜ最初に物理分野でうまくいったのか、そのうまくいった要素は他に展開できるのか。私の分野の化学で、生長さんの話を聞いて一つ作業仮説を立てるならば、そのコミュニティのトップジャーナルのエディターが受容するかどうかということがあり得る気がします。そのあたりはパネラーの皆さんの領域でどうですか。トープジャーナル

のエディターが「プレプリント、オーケー」と言うと、傘下のジャーナルみんな横並びになりそうな気がして、それが物理で起きていて、化学ではまだ起きていないということだと思うのですが、いかがでしょうか。

●**生長**　全くおっしゃるとおりだと私も思います。端的に言うと、日本化学会がプレプリント投稿を受け付けて、化学会全体でこの流れを進めようと言って、評価基準をみんなで修正していこうという流れをつくれば、時間がかかるかもしれませんが、徐々にはそうなっていくのではないかと思います。その過程で、プレプリントは有益なものであるということがコミュニティに受け入れられていけばそれで問題はないかと思います。でも、誰がそれをやるのだというのが難しいところだとみています。

●**フロア 4**　お隣に物理学会の関係者もいらっしゃいますよ。

●**フロア 8**　物理の中でも、先頭を走ったのはやはり高エネルギー物理なのです。ビッグサイエンスで、データは共有物という意識もあったと思いますし、何といってもコミュニティとしては小さいので、そこでは合意が得やすかったということがあります。次第に物性物理にも広がっていったのですが、私の見るところ、その契機は、高温超伝導ではなかったかと思うのです。あのときに研究のスピードが爆発的で、査読付きの論文なんか待っていられないと、みんなプレプリントで勝負したのです。それで一気に広がったというのが私の個人的な印象です。

●**フロア 4**　これは NII の北本さんが以前ディープラーニングで紹介した例とまさに同じです。研究が早過ぎて待っていられない。

●**坊農**　北本さん、何かコメントはありますか。

●**フロア 9**　NII の北本です。昨年の第 3 回 SPARC Japan セミナーで話したのは、ディープラーニングで研究の速度が爆発的に速くなって、論文の査読を待っていられないからみんなプレプリントに流れたということです。高温超伝導と全く同じというのは今日初めて知ったのですが、それは非常に大きな要因かと思います。

●**坊農**　どうもありがとうございます。昨年出ていない人にもお分かりいただけたかと思います。

今の、化学ではこれから起こるとか、そういう話を聞いていて、われわれ生命科学の分野はそういう意味では不思議な立場だと思いました。ビッグラボなどが何かやってというところが日本ではあまり起きていなくて、海外で分野の最先端を行っているところが最新の機器を使ったデータなどをプレプリントで出しているということは、たまに聞きます。日本では生命科学関係ではそういうことがあまり起きていない感じがします。化学などは、まだ学会がきちんとオーガナイズする力を持っているのではないかと思いました。やはり日本化学会に頑張ってもらって、プレプリントをやりましょうと言うのが一番いいのでしょうか。

●**生長**　プレプリントの普及の順番として、先に物理があって、次に生物が来て、化学に行くという順番は分からなくもないです。実験に時間がかかり過ぎて、査読の速度とキャリアプランに齟齬が起きやすいのが生物学の世界だと思っています。生物学の研究を少ししているので、最近何となく感覚が分かるようになってきました。先ほど言われていた高温超伝導の話は、研究のスピードの方があまりに速過ぎてということですが、評価やキャリアパスなどの人間的なところとサイエンスのスピードのギャップが出てきた瞬間に、プレプリントが普及していくのではないかと思います。

化学の世界では、今までは評価やキャリアパスなどのサイクルと研究サイクルがちょうど良くかみ合っていたので、プレプリントの必要性がひょっとしたらな

かったのかもしれません。けれども、化学の世界でも生物との融合的な研究をかなりたくさんの人がやるようになった結果として、プレプリントが必要という話になったという気がしています。それをトップの方が理解してそれをやろうとするかなとは思います。

●**フロア 7**　北本さんの補足ですが、コンピュータサイエンスのプレプリントが普及したのには、もともとジャーナルに権威がないという事情もあります。カンファレンスがベースで、カンファレンスにも権威があるもの、ないものがあって雑多な世界です。しかも昔からテクニカルレポートも大学ごとに出すという文化もあって、そういうことに関してもともと無頓着なところがあったのです。今までも、例えばページランクの詳細な論文はスタンフォードのテクニカルレポートしかないということがあったぐらいで、それがたまたまアーカイブに載るようになったのが割と最近です。コンピュータサイエンスの場合は逆にジャーナルにさほど権威がなかったのでそういう方向に流れてもそんなに摩擦はなかったと私は理解しています。

●**坊農**　確かに考えてみると、私が生命科学でプレプリントでと例に出したのも次世代シーケンサーで、より長く早く安く読もうという競争があって、早く論文を出したがっている分野です。小野さんに出していただいた例も、シングルセル、1 細胞の技術という、今競争の非常に激しいところです。生命科学の bioRxiv を見ていても、そういうところしか出ていないような印象を受けています。厳しい競争がプレプリントを促すというまとめでよろしいのでしょうか。そういう競争が化学でもいろいろありそうですが、化学だと次は何がいけそうですか。

●**生長**　私が属している有機合成は研究サイクルが短い分野なので、タッチの差は当然あるのですが、プレプリントで評価される、されないというほどの差ではない気がしています。どちらかというと化学はもっと研究スパンの長いもので評価が決まってくるようなものが増えてきたと思います。

●**坊農**　それでは時間となりましたので、全体議論を終了いたします。先生方、どうもありがとうございました。

SPARC Japan NewsLetter
ISSN 1883-826X
NO.35 2018 年 12 月

SPARC Japan ニュースレターでは、各回セミナーの報告に講演やパネルディスカッションを書き起こしたドキュメントを加え、さらにそのほかの SPARC Japan の活動をご紹介します。

CONTENTS

■ SPARC Japan 活動報告
SPARC Japan 運営委員会
国立情報学研究所実務研修（CERN）

■ SPARC Japan セミナー報告
企画概要
参加者から
企画後記
ドキュメント
（講演・ショートディスカッション・パネルディスカッション）

■ SPARC Japan 活動報告

CC BY

SPARC Japan 運営委員会

SPARC Japan 運営委員会の会議資料をウェブサイトで公開しています。
http://www.nii.ac.jp/sparc/about/committee/

国立情報学研究所実務研修（CERN）

国内外の OA イニシアティブや関係組織と連携し、学術情報流通の更なる発展に取り組むという基本方針のもとに、2015 年度に高エネルギー物理学分野の情報サービスに係る国際連携協定を欧州原子核研究機構（CERN），高エネルギー加速器研究機構（KEK）及び国立情報学研究所（NII）の 3 機関で締結しました。その具体的活動として国立情報学研究所実務研修生を CERN に派遣しました。

研修期間：平成 29 年 9 月 30 日（土）～平成 29 年 12 月 28 日（木）
所属機関・氏名：東京大学附属図書館・横井 慶子
研修場所：欧州原子核研究機構（CERN）
研修テーマ：物理学分野のデータベースに関するデータキュレーション
https://www.nii.ac.jp/hrd/ja/jitsumu/h29/index1.html

■ SPARC Japan セミナー報告

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

2018 年 2 月 21 日（水）　国立情報学研究所 12F 会議室　参加者：67 名

本セミナーは、私たちが向かうべき具体的活動の方向性を明らかにするために，科学と学術の本来の姿を議論することによって，学術知識の生産に従事する関係者にとっての本筋を見出し，参加者が自らの状況に照らして次に進む一助となることを期待して企画されました。

次ページ以降に、当日参加者のコメント（抜粋）、企画後記およびドキュメント全文（再掲）を掲載しています。その他の情報は SPARC Japan の Web サイトをご覧ください。(https://www.nii.ac.jp/sparc/event/2017/20180221.html)

企画概要

2002 年のブタペスト・オープンアクセス・イニシアチブと 2013 年の G8 オープンデータ憲章は，科学アカデミーからのボトムアップな自主活動とトップダウンの政策提言という互いに交錯しつつも共通の方向を目指し，「オープンサイエンス」の実現を加速する転回点となった。

パネルディスカッション（左から David 氏，深貝氏，村山氏，Joseph 氏，倉田氏，市古氏）

この変化の中で，我々が向かうべき具体的活動の方向性を明らかにするためには，このセミナーの一日だけでも立ち止まり科学と学術の本来の姿を議論することによって，学術知識の生産に従事する関係者にとっての本筋を見出し，参加者が自らの状況に照らして次に進む一助となることを期待している。

しかし，この方向が社会全般にわたるデジタル化の動向と軌を一にするものであるとしても，人類がこれまで営んできた科学という活動の本来のあり方をついに実現しようというのであるのか，それとも，デジタル化といういわば道具の変化が科学活動の本質を変えようとしているものであるかの見極めはついていない。実際，デジタル化とともに学術コミュニケーションの形態は大きく変化し，研究室，教室，「図書館」における知識の生産と伝達の様式は予測不可能なほどに変貌しつつある。

Paul A. David 氏

参加者から

（大学/図書館関係）
・今後、研究を考えていくうえでオープンサイエンスに向き合う必要があると感じました。ただ、研究大学を目指さない大学では、その必要性を感じている人は少ないため、大学のトップの指示がない限り対応しないままになってしまうような気がしました。

（企業/図書館関係）
・最後のディスカッション、面白かったです。まさに"オープンサイエンスを超えて"いたと思います。ただ、図書館に関連する話ももう少し聞けたら良かったかと思います。

（大学/研究者）
・素晴らしい企画でした。SPARC の普段の企画は少し実践よりのように感じていましたが（それ自体の意義は理解していますが）、そろそろ、こういう「科学や知」に焦点をあてた「高い視点（ワイドスコープ）」の企画や議論があってもいいように思いました。時代が時代なので。本当に面白かったです。〔動画中継視聴者〕

（その他/研究者）
・紙ベース、出版ベースの科学が、デジタルベースの科学に変わっていく中で、“オープンサイエンス”をどう捉えるべきなのか、もっと考えるべきと感じました。科学の民主化は 1 つのとらえ方なのですが、それが具体的に何なのか、もっと多くの人が考えるべきかと。研究成果管理システムは、民間サービスでいいのではないかと思うところ。図書館のあり方を考えるとき、利用というものがデジタル化されたことによって大きく変わっていることも注意すべきと感じました。「モノ」が存在した時代と、デジタルで存在している時代で異なるとも思います。

企画後記

CC BY

企画当初 9 月、運営委員から「オープンサイエンス」という非常に抽象度の高いお題と参考文献が提示され、文献を読みながら普段の業務における具体的思考レベルから抽象的思考レベルに一気に引き上げられる衝撃を受けました。企画では、米国 SPARC と協調し、「オープンサイエンス」を深く考えることを心がけ、登壇者を決めました。3 年間の企画活動を通して最も充実しましたが同時に最後まで登壇者が決まらない最もハラハラした運営でもありました。このような勉強の機会を作っていただいた関係の方々、共に働いてくださった方々に感謝したいです。

蔵川　圭
（国立情報学研究所）

論文から研究データへ、オープンアクセスからオープンサイエンスへ議論が拡張される中、そもそもサイエンスとは何かというのは本来避けて通れない議論です。これに正面から向き合うことができたのは、これまで SPARC Japan の積み重ねがあったからこそと信じています。本セミナーの企画がまとまるまでの担当者同士の対話は、それぞれの知識と経験が融合した非常にエキサイティングなものであったことをここに記しておきたいと思います。

林　和弘
（科学技術・学術政策研究所）

今年度第 2 回セミナーに引き続き企画に参画させて頂きました。オープンサイエンスと図書館の関りは、未だは明確な回答はないのかもしれませんが、図書館職員が自らの価値を高め、研究支援に本当に取り組むのであれば、避けては通れないものであると考えます。SPARC Japan セミナーの企画 WG は 3 年間務めさせて頂きました。今後も本セミナーが学術情報流通と研究者と図書館職員をつなぐ一役を担うことを期待します。

梶原　茂寿
（室蘭工業大学附属図書館）

Heather Joseph 氏

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

趣旨説明

蔵川 圭
（国立情報学研究所）

蔵川 圭
http://researchmap.jp/kurakawa/

今回のセミナーは、「オープンサイエンスを超えて」と題して企画しました。今、学術情報流通の世界では、オープンアクセスからオープンサイエンスへと主要なスローガンがシフトしてきています。SPARC Japan セミナーにおいても、私が企画に関わりはじめた 3 年前から、研究データ共有、オープンサイエンスについて採り上げ、今回でシリーズ 6 企画目となります。

本日の課題

シリーズを通して、さまざまな関係者にご登壇いただき、研究データ共有、オープンサイエンスに関わる活動について学びました。そして、なぜ今オープンサイエンスという言葉が声高に叫ばれているのか、オープンサイエンスという概念は一体何で、どこから生まれ、どのように波及して、結果としてわれわれの仕事の問題意識の中に浮かび上がってきているのか。また、高度にデジタル化されつつある学術コミュニケーションの関係はどうであるのか、知識生産の方法は革新的に変化しているのか、本日はこれらの疑問を解き明かしたいと考えています。オープンサイエンスという概念の起源と、現代の学術情報流通における先進的な取り組みについて、世界で活躍する一流の講師陣をお招きしてご講演いただき、今日一日皆さまと共に考えてみます。

学術コミュニケーション形態の変化の中で

過去を振り返ると、2002 年のブダペスト・オープンアクセス・イニシアチブの宣言と、2013 年の G8 オープンデータ憲章の制定という二つの出来事は、科学アカデミーからのボトムアップな自主活動とトップダウンの政策提言という互いに交錯しつつも共通の方向を目指したオープンサイエンスの実現を加速する展開点となりました。この方向がわれわれの社会を取り巻く、いわゆるデジタル化の動向と同じであるとしても、人類がこれまで営んできた科学という活動の本来のあり方をついに実現しようというものであるのか、それともデジタル化といういわば道具の変化が科学活動の本質を変えようというものであるのかという見極めまではついていません。

確かにデジタル化とともに学術コミュニケーション

の形態は大きく変化して、研究室、教室、図書館における知識の生産と伝達の様式は予測不可能なほどに変貌しつつあります。このような変化の中で、学術知識の生産に従事するわれわれの日々の活動の方向性を見いだすために、今日一日だけでも立ち止まって、科学と学術の本来の姿にさかのぼり、現代の知識生産のあり方とその環境について議論してみたいと思います。皆さま方の仕事に必ずやお役に立てると確信しています。

本セミナーは、科学技術・学術政策研究所の林和弘さん、室蘭工業大学附属図書館の梶原茂寿さん、横浜国立大学大学院の深貝先生と共に企画いたしました。本日は一日どうぞよろしくお願い申し上げます。

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

オープンサイエンスを真に理解する：

その有益性の潜在能力、脆弱性、機能的パフォーマンスの問題、これらの解決策を講じない方法

Paul A. David

(Stanford University)

CC BY NC ND

講演要旨

講演では「オープンサイエンス」の正しい意味について三つの関連する要点について述べる。それは持続可能な経済成長のための潜在的機能、この現代の社会制度の不本意なパフォーマンスの本質と原因、それにこれらの問題の改善に取り組むことを避ける方法である。これらの了解事項のもとに、科学技術政策についての現代の議論や決定に必要な基準を提供する。

第一に、「オープンサイエンス」は機能的に区別されたサブコミュニティの行動的相互作用を伴う多次元動的プロセスとして最もよく理解されている。ここでいうサブコミュニティとは、教育者、理論・実証研究者、読者、執筆者、評論家、研究資金提供公共機関と民間企業組織、出版者、公文書保管者、雑誌・書籍編集者、および査読者を指す。これらのサブコミュニティは、それぞれがローカルであっても国際的であっても、外部から押し付けられたりすることも、何ら独立した執行力もない関連規範的構造を有する。サブコミュニティの範囲内で、あるいは少なくともその国際的な範囲内で最も一般的になることで、「オープンサイエンスの規範」は最も十分に説明され、親しまれるものとなってきている。従って、機能的なパフォーマンスが個々の規範を遵守する可能性およびより広く理解され評価されるべき規範順守による相互作用の結果を簡単に見直すことから始める必要がある。

第二に、「オープンサイエンス」のプロセスは、科学資源がその運用規範に沿った方法によりミクロレベルで配分されるようにマクロ動学レベルで機能し、また競争の厳しい市場ベースの資源分配効率において有益なミクロレベルでの仕組みを補完する配分効果を生み出すことも期待できる。結びついたサブシステムが引き起こす可能性のある（地球温暖化など）「負の外部性」による弊害がないのであれば、持続可能な経済成長をもたらすためにサブシステムの相互依存的行為に依存することができる。これを理解することは、科学技術政策戦略と手段的戦術の選択が寄るべき基盤となる。

第三に、すべての人間社会のシステムのように、規範がはっきりと示されているところではどこでも、従って研究室や企業のオフィスでも個人の逸脱行為は予想できる。同様に、制度および組織設計の失敗は、これらの法社会組織体の指導者たちが広める規範となる指針を歪めてしまう。従って、指導者たちは、一方ではコストのかかる市場での失敗に対処するための、また他方では、相互に等しく権限をもった上での信頼の礎を損なわないように、またサブシステムの知識を生成し、普及させる機関の内部是正能力を超えないように個人の科学的不正行為の広がりを抑えるための継続的な是正措置を求めている。オープンサイエンスの構造的機能は歴史的に発展してきており、その過去の制度化されたレガシーの存続は、現代の機能不全の結果の発生源である可能性がある。しかし、後に制度によるレガシーは、特に教育科学研究活動において「フリーオープンソースソフトウェア」を導入し成功して一般的に認められたことで、レガシーを時代遅れのものとし、また科学的活動のあらゆる面でソフトウェアを実装した「開示性」により容易に置き換えられるようになったという曖昧な理由で早急に処分されるべきではない。オープンサイエンス・プロセスの必然による人間の社会性と、アルゴリズム情報プロセスの限界を理解することで、私たちがオープンサイエンス・システムの仕組みにおいて直面する厄介なパフォーマンスの問題は、伝達、知識を共有する人的要因のための「オープン」コンピューター・アルゴリズムの統合型アレイを取り換えて容易に問題を「解決」できる性質のものではない、という誤った考えを速やかに払拭する。

Paul A. David

スタンフォード大学経済学名誉教授、同経済政策研究所シニアフェロー。オックスフォード大学経済学・経済史名誉教授、同オールソウルズカレッジ名誉フェロー、オックスフォードインターネット研究所シニアフェロー。これまでに150件以上の雑誌論文を発表し、雑誌編集にも寄与、Technical Choice、 Innovation and Economic Growth (1975)、The Economic Future in Historical Perspective (2003)を含む複数の図書の著者であり編者である。数量経済史の先駆者のひとりであり、アメリカ経済史、経済歴史人口統計学、科学技術経済学の分野において幅広く貢献してきたことで世界的に知られている。

主な研究テーマは経路依存性（ミクロ・マクロ経済現象における歴史的な出来事の持続的な影響）を生じさせる条件の調査研究である。現代経済政策研究の２つの主な領域は過去20年の彼の成果に置いて明らかになっている。一つは情報技術の標準化とネット業界の進展、もう一つは公共部門における科学研究の資金援助と管理上の法的機関と社会規範の影響、公共と民間部門のR&Dの関係である。最近はフリーオープンソースソフトウェアの組織、性能、実行可能性についての国際研究プロジェクトを率いている。

本日この場にいられることを非常にうれしく思っています。そして、国立情報学研究所とその優秀なスタッフの皆さん、蔵川圭さん、林さんなど、本日私がお話しし、皆さんにお伝えしたいと思っている内容について、私がより理解を深めるための準備として、数日にわたって非常に刺激的な会話をさせていただいた方々に感謝申し上げます。

私の講演のタイトルおよびサブタイトルは、今回の私の講演の意図を大方表現しています。つまり、「オープンサイエンス」という言葉に対する私の理解や、先にお話ししてくださった皆さん、またこれから講演される方々が、オープンサイエンスについてどのように理解されているかということです。人間の福祉と幸福、および知識の向上が原動力となる経済成長という両方の観点から、知識の向上、より深い知識を得ることによって恩恵を受けられるという理由で、いかにオープンサイエンスが重要なのかについてお話しするつもりです。また、社会経済的知識システムとしてのオープンサイエンスについてもお話ししたいと考えています。それは、一部の非常に重要な側面においては、制度化された形を取っています。しかし、オープンサイエンスは市場によって直接的に導かれるわけではないという点においては、脆弱なシステムです。販売される商品を生み出していないからです。知識は、オープンサイエンスシステムにおける私的市場システムの財産として扱われるべきではありません。そのシステムについて理解したら、次にそこに生じる問題やシステムの機能的遂行についてお話ししたいと思います。完璧なシステムは存在しないに等しく、他と比べてより深刻な問題もあります。それはつまり、システムを見守る人々やその一部になっている人々が、そのシステムは恐らく改善・修正可能であろうと考えていることを意味しています。私はシステムの修正および改善策の提供方法について皆さんに話すつもりはありません。逆に、問題を解決しない方法について話したいと思います。

システムとしてのオープンサイエンス

アウトラインやアジェンダの代わりに、パート 1、2、3 に対応する三つのコースのメニューを提示します。最初の重要なパートはオープンサイエンスの理解です。オープンサイエンスは、世界的な情勢ではありません。システムは確かに存在していますが、動的なものです。変化しており、進化を遂げています。特に、現代社会においては独特のシステムです。市場への浸透が組織を動かし、インセンティブは人々のサービスの市場価値およびお金を使って他の物を買う力に結び付いています。このシステムの中では、市場取引は人々をやる気にさせたり、人々の貢献度合いを評価したりする中心にはありません。だからこそオープンサ

イエンスは独特であり、なおかつ極めて重要です。過去から進化を遂げてきたもので、深い歴史的ルーツを持っています。システムの歴史的遺産について言えば、それは前の時代からの贈り物です。西側においてそれは社会の封建的組織によって特徴付けられたものです。これは、マルクスが上位有産階級や彼らのために働くプロレタリアートに用いたものではなく、歴史家のマルク・ブロックが特にヨーロッパ的だと考えたシステムですが、ヨーロッパのどこにでもあるというわけではありません。それは、土地資源の支配権を主張する強い人々がいて、彼らにとって（第一に軍事的意味で）有益な部下に、土地の特定の部分の資源を利用する権利を与えるという隷属的システムです。従ってこれらの人々は土地を手に入れます。このため、彼らは土地をくれた人にサービスを提供しなくてはなりません。この取り決めは、ヨーロッパの前封建的時代に遡る非常に古いものです。このシステムが支配階級間での関係を築き、より近代になるにつれ、その支配層が自分たちの宮廷に、さまざまな種類の特別な知識を持っていて、誰によっても成し得ないサービスを提供する人々を求めるようになりました。「feudalism（封建制度）」という語の語源は、古代フランス語の「*féodal*」です。それは、「fief（封土）」に関する事柄に関係があったに違いありません。封土とは、特別なサービスを提供できる者の土地のことでした。始まりは特定の兵役で、その後、書記官に進化しました。土地の領主のためにメッセージを書いたり送ったりできる人々のことです。当時の領主たちは必ずしも読み書きができたわけではなく、実際のところ、読み書きができた領主はほとんどいませんでした。

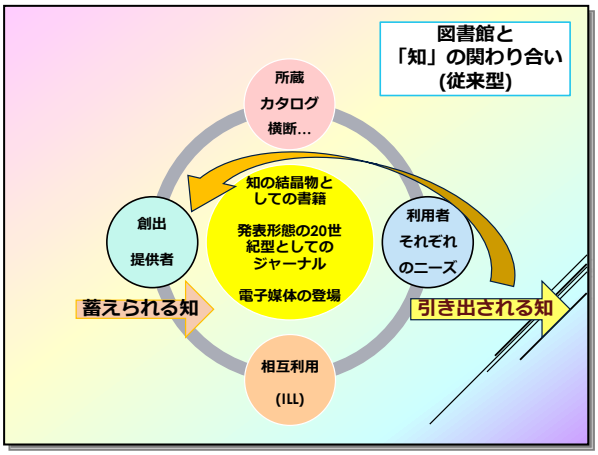
従って、このシステムは、アラビア数学とインド・ヨーロッパ数学の混合を通じて知識領域と絡み合っていました。こうしたものは、宇宙や惑星、太陽系、われわれの暮らす地球上で見られる現象といった物理的システムを調査するための強力なツールを与えました。数学と連動する知識領域がこのシステムを形成したのは、その利用が科学的組織および科学的コミュニケーションにとって有効だったからです。それは 17 世紀科学革命と絡み合い、重要な要素となりました。ここで言う 17 世紀科学革命は、科学革命とは同じではありません。それはネットワークを生み出す組織形態で、それを通じて当時の新しい科学的発見に関する情報が伝播し、さらなる進展のための基礎となり得るのです。それらはこの社会システムの特徴ですが、今回私が皆さんに印象づけたいこととは別の問題であり、これが他の多くの社会システムと同様に、システムの複雑なところなのです。

パート 2 では、オープンサイエンスシステムにおける三つの重要なパフォーマンスの問題について話します。それは三つのレベルで発生します。一つ目は、個人行動のレベル。二つ目は、組織内での整合性と、組織のパフォーマンスに影響を与え悪化させる、組織内での論争の欠如というレベル。三つ目は、知識の別の側面とその適用に、科学的資源の誤った分配によって悪影響がおよぶというレベルです。例えば、科学的資源が 1 カ所に集中し過ぎるという方向性の間違いや、システムを悪化させがちな構造上の問題です。

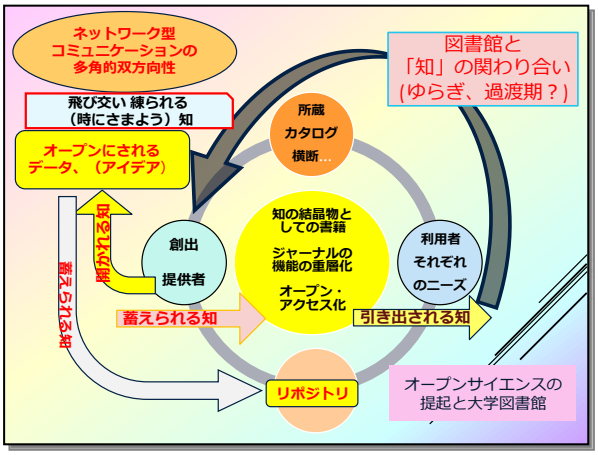
パート 3 は非常に簡潔です。なぜなら、こうした問題全ての解決策を私は持ち合わせているわけではないからです。今まで話してきた 3 点について私は熟考してきました。そして、全ての場合に当てはまる一つの政策的なメッセージを考えました。それは、医者が医療の世界に入るときに従うことが求められるメッセージ「害をおよぼさないように」ということです。つまり患者を診察するときに最も重要なのは、患者を死なせたり、障害を残したり、患者の命を救うために不可欠とは限らない方法で患者を傷つけたりしないということです。従って、私は自分が示した三つのタイプの問題の解決を妨げるやり方について、警告的メッセージを幾つか提示するつもりです。

オープンサイエンスモードの豊富なリポジトリの役割

本日お集まりの皆さんの多くが、従来型の図書館やデジタル資料のリポジトリで働いており、そうした資料のキュレーションや検索法に取り組んでいるとお聞きしました。つまり、皆さんは、特定のトピックの検索によって入手可能となるように、いかに内部情報を構築できるかということに取り組んでいます。また、私が話そうとしている枠組みがご出席の皆さんには認められないかもしれないというためになる洞察を頂きました。しかし、中心的な問題は、2 年前に第 18 回図書館総合展で発表された深貝教授の「科学的知識のオープンな基盤：リポジトリを介した知のやりとり、そして、大学図書館の果たし得る役割」のスライドの最初にある素晴らしい図に表されています（図 1・2)。ポイントは、これが現在移行中のシステムであるということです。それは複雑なシステムです。当事者個人や当事者グループの活動と、皆さんが行ける図書館の役割を持つ機関や、デジタル形式であるために必要な情報を抽出できるリポジトリとをつなぐシステムです。抽出は非常に容易で、高度に的を絞り、精選することが可能です。



（図 1）



（図 2）

知識の生産における図書館の役割は伝統であり、図 1 に示されています。それはニーズに直結しています。それぞれのユーザーにそれぞれの研究上のニーズがあります。過去の研究資料など、研究に必要な資源を利用するといったニーズです。研究者が作成したデータは今や保存可能です。皆さんはそのデータにアクセスして、自らの目的のためにそれを利用することができます。従って、こうしたデータが体系的に保存されていることに対するニーズがあるのです。そのモデルでは、原則として、保存および検索の機能を提供するのは図書館員です。その知識は研究者のニーズの一部を満たし、良い循環が生まれます。

図 2 は、皆さんが取り組んでいるシステムの図です。進化し、新たな機能が加わっています。デジタル情報技術の発展やデジタル化資料でできることの発見によって生まれた機能もあれば、非常に新しく、使いこなすためには、リポジトリを提供する人々の適応が必要な機能もあります。従って、われわれはまだこうした技術の利用から得られること全てを見つけたわけではなく、その活用、生成、伝達といった方法も全てを見つけたわけではありません。

デジタル資源に自由にアクセスできるようにし、資源の配布や使用、キュレーションの方法や利用記録、資源使用の特許権といったものを構築したら、資源は世間一般の人々にとって使いやすいものになり得ます。リポジトリには非常に中心的な役割があり、リポジトリに保存されたコレクションは、オープンサイエンスモードにおける資源分配および研究活動に関わってくるのです。

オープンサイエンスを理解する

このシステムを理解するには、共通の規範に基づいた構造を導入することで重要な補強を求めるというアイデアから始めるのがより簡単です。規範というのは、ある社会システムの中で他者と共生する際に、何が相応で適切な行動かを教えてくれるものです。もしあなたが、今とは違う文化の一員になる場合、どんな風にふるまうべきか、その文化の中の人に教えてもらわなくてはならないこともあるでしょう。これが行動的規範です。行動的規範は制約につながることもあれば、このシステム内で働く代々の人々の業務モードを組織、強化し、伝達する必要がある正式な機関の機能を向上させるという発展につながることもあります。

私は皆さんに「オープンサイエンス」と呼ばれる変わったシステムがあるとお話ししました。これは市場システムではありません。市場システムは別に存在しています。企業は各自の研究所で研究を行い、その発見の特許を取りライセンス化します。そして製品を商品化し、ユーザーと相互交流をします。ユーザー学習は、新製品および新サービスの生産者にとっての学習と理解の源であり、研究を支えるものです。

われわれはなぜオープンサイエンスと商業化サイエンスという、二つの全く異なるシステムを持っているのでしょうか？　この二つのシステムは、さまざまな目標を持つ個人および組織の行動に影響をおよぼすさまざまなモチベーションを生み出すため、多くの点で対立します。事業会社または新興企業の純資産増加という目標を持っている一方で、財産の移転に関係するために、あるいは持っているものがアイデアや知識であるために商品販売では手に入れられない社会資源を、研究者たちにキャリアを積ませて集められるようにしたいという目標も持っているといったことです。知識はお金と引き換えに取り引きするのが非常に難しいものです。私がある人に「清涼飲料向けの素晴らしいアイデアがあります。たっぷり稼がせてあげますよ」と言ったとします。するとその人は「なるほど、それは何ですか？」と言うでしょう。私は、「もし私がそのアイデアが何か言ったら、あなたはそれを手に入れてしまいます。アイデアを手放さないであなたにどうやって伝えたらいいのでしょう？」と言います。そのため、われわれは新しい知識の取り引きに複雑な方法を取ります。知識を受け渡すことで、どのように報酬を得られるのか？　これがアイデアの出発点であり、研究者は知識領域でやりたいことを続けるために何らかの利益を得るためには、アイデアを譲り渡し、公開し、他者が、自分たちの研究が不正ではないこと、自分たちの知識がなければできないことがあるということを検証するのを認めなくてはなりませんでした。

なぜ現代社会のほとんどが、互いに矛盾するように思われるシステムを持っているのでしょうか？　その理由は、それらはまた互いを補完してもいるからです。この二つのシステムが適切なバランスで共に維持されれば、どちらのサブシステムも一つだけでは不可能なものを生み出します。そのようなシステムをわれわれはどうやって手に入れたのでしょうか？　われわれは、これを理解する必要があります。新システムのルーツは、システムの中心的な目的、また、なぜ人々がそのようなシステムに固執するのかをしばしば明らかにするからです。

オープンサイエンスの歴史

オープンサイエンスの不可解な現象について詳しく述べたいと思います。知識のように、危険で破壊的なものを秘密にすることを特徴とする世界において、どのように人々は知識を発表し、他人がそれを試したり、うまくいくかどうかを調べたりできるように知識を誰もが入手・理解可能なものにするようになったのでしょうか？

社会科学である現象が見られるとき、あなたは根本的に異なる二つの方法でそれを説明しようとすることができます。一つは、それがどのように発生したかを見いだし、その論理的起源まで広げることです。それが論理的に存在していなければ、人々はそれをつくりたいとは思わないでしょう。そのため、人々は何が欠

けているのかを理解し、こう言わなくてはならないでしょう。「これが、『なぜわれわれにこれができないのか、なぜわれわれがそれを持っていないのか』という問いへの論理的回答です」。それで、本質的な機能的説明を得るために、人々は「このシステムのプロセスはわれわれに何をするのか？　その機能は何か？」と尋ねます。機能がなくては、人々はそれが娯楽か象徴的な目的のためでない限り、何かをし続けることはないでしょう。しかし機能的レベルでは、他のものを変換してエネルギーを生み出し、それをコントロールし、物理的な商品を巧みに扱い、説明が必要な新しいものや、これらの新しいものがいかに適用されるかといったことを発見しようとすることによって、知識が世界で力を発揮するニーズがあります。それらは機能的な理由です。しかしながら、論理的な起源は、「われわれはこれが必要で、それを今日発明するのだ」と人々が言うことに基づいています。それが必要で、論理的にそのためにそれをつくった時期が過去にはあったに違いありません。

ですが、これは事が始まらなかった理由の説明にはなっていません。最初から見てみましょう。最初に人が「何かしらこういうものが必要だ」と言ったときから、なぜ何世紀も過ぎてしまったのでしょうか。なぜ人々は反対のシステム内に暮らしていたのでしょうか。それはつまり、彼らがその物を共有する必要がなく、協力する必要も、できるだけ多くの人がそれを理解できるようにそのものを入手可能にする必要も、それを理解できるように人々を教育する必要もなかったということを意味しています。16 世紀以降、ヨーロッパにおいて、われわれは隷属という支配的なシステムの中で人々を動機づけるのに十分な条件を設けたでしょうか？　資源を豊富に持つ人はいました。貴族の地位の高い者も低い者もいました。宮廷内や所有地で披露するものを必要としていた人々もいました。彼らは兵器や防御について知っている人々の知識を必要としていました。彼らは、矢や大砲の砲弾などの軌道を計算する必要がありました。火薬が発明されてからは、大砲の砲弾その他の投射物に耐えられる城を設計可能にする必要がありました。

パトロネージュは古くからありましたが、そこに新しい数学が問題を引き起こしました。貴族たちは自らの宮廷に、数学的要素の大きい新しい知識について知っている者を置くことに興味を持っていました。これらの者たちが数学的にやっていることを理解できる貴族はほとんどいませんでした。そのため、こうした人材は欲しいが、その人材が優れた博識な人物なのかどうかが分からないという問題がありました。例えば、素晴らしいリュートをつくることができるという者がいれば、あなたはその作品を見てそれが本当かどうかを自分で確かめられます。しかし、貴族たちは、数学的知識のある人物が本当に博識かどうかを実際に自分で判断することはできなかったのです。貴族は、自分では理解できない専門知識を持つ人々を必要としたけれど、その能力を理解するためには自身も専門家にならなくてはならなかったということです。

これは「情報の非対称性」と呼ばれています。それは現在でも政府契約の大きな問題として存在しています。例えば、請負業者は新しい輸送システムや現代的な空港デザイン、新しい軍備システムを構築できると言うことはできますが、それをどうやって信じたらいいのでしょうか？　これは全ての専門的サービス分野に存在する問題です。あなたの主治医は事情が分かっていますから、あなたに状況を説明しようとすることはできますが、あなたが医師としての訓練を受けていなければ、主治医は説明するのに大変多くの時間が必要でしょう。あなたは医師の話が理解できなければ、その話が真実かどうか分かりません。従って、あなたに必要になるのは人を信頼する根拠であり、正しさを認証する何らかの方法です。あなたはその人物には会ったことがない人々にこう言ってもらう必要があります。「彼または彼女の評判は彼女を上回っています。彼女はその分野で非常に有名です。確かにこの人は信頼すべき人物です。　その人があなたに何かをやれと言えば、あなたはそれを受け入れるべきです。同業者

から高い評価を得ている人物だからです」。また、話題になっていることを知っていて、「はい、この人がした仕事は知っています。私がそれを信用するのは、実際に見たことがあり、それが機能しているのをこの目で見たことがあるからです」あるいは「知り合いがそれを試したことがあります」と言ってくれる人を抱えることもできます。従って、人々がこの情報の非対称性を解決する知識があるとはっきり示せることが重要になりました。

制度化されたオープンサイエンスの規範的構造

制度化された規範の把握に関しては、1938 年に初めてロバート・K・マートンが現代科学を実行する正しい方法について記しています。彼は 1970 年代に本を編纂する前に、それを自分の論文と記事に入れて発表したいと考えました。実は彼は 1942 年の第 2 次世界大戦中にある論文を発表しましたが、そこには強烈な政治的敵意がありました。ある意味それは、ナチスの科学と、19 世紀初期にドイツの物理科学と化学、物理学を先導する立場にあると一般的に認められていたものの破壊に対する攻撃または批判でした。彼は独裁主義的なコントロールではなく、協力を強調しました。協力は自発的なものです。科学を実践できる人については、出生地や文化、宗教に基づいて何らかの帰属的な制限はあるものの、普遍主義とは科学が全ての人に開かれていることを意味しています。科学への参入は明らかな価値があるかどうかに基づいてきましたが、人々は特殊な意味で、知識が自分たちに個人的に利益を与えるかどうかには関心を持つべきではありません。すなわち、科学を実践する人々は真実に興味を持つべきで、その興味は、個人的なことには無関心な情熱的なものであるべきです。彼らは問題とその解決策を見つけることに興味を抱きます。

開放性には二つ意味があります。一つは探求の開放性で、もう一つが懐疑的態度への開放性です。つまり、自由に誰かに挑戦し、積極的に懐疑主義を実践する能力です。これは、「ここまではあなたの言うことを理解できましたが、これはどうなのですか？　この部分は論理的に理解できないと思われます」と言う能力です。言い換えれば、それは質問する自由です。それが理想的な社会化された規範を生み出します。研究成果および方法の完全な開示は、同業者の複製による検証への期待と同じく、必要なことの一部です。それらが条件です。それらが、質問し、証拠を見たいと言う人々によって、個人的には感情を害されないという社会的規範です。

手続き上の取り決めは、あなたの評判が高い場合、同輩が「そうです。この人は非常に優れた仕事をします」と言う場合に起こることに基づく報酬次第です。それは、あなたに大学で教えさせ、あなたに大学で研究させようとする大学の報酬を増やす傾向があります。同様にそれはあなたに助言を求める企業にもつながるかもしれません。必ずしもあなたを雇用する必要はありません。何らかの方法であなたの知識を求めるのです。評判は、あなたの政治的なかけひきその他ではなく、同業者によるあなたの科学的貢献の評価に基づきます。評価に値するかどうかは、著者の非帰属的な特徴と、発見または発明を優先してほしいという当然の主張を含むその貢献次第になります。最後に持ち上がる問題は、誰かが働いているかどうかをどのように判断するかということです。あなたはその人たちが何をしているのか監視することはできません。温かいお風呂に浸かっている物理学者が、問題について非常に熱心に考えていたとしても、お風呂にいる彼らを見るだけでは、見た目には楽しそうに見えるだけで、彼らが熱心に考えていると知ることはできません。

できるのは、誰かがそれを最初に発見したと主張し、それを最初に示したのであれば、その人物に何かを与えることです。従って、開示することが、自分が誰か他の人がやったことをコピーしているのではなく、自分でやったと主張する重要な方法になります。主張の優先性が重要になります。それは、時に優先性の対立および主張と共に議論になります。これは、このシステム内に存在する評判を巡るライバル関係という側面

であり、それが非協力的な行動を生み出します。あなたはライバルを遅らせたいと思っていますが、一方で、あなたは自分の取り組んでいることを開示しなくてはならないでしょう。あなたはライバルより先にそれを発表したいと思っているので、システム内の内面的緊張という側面もあります。

そのシステムの科学革命との関係については、リーダーたちは、科学的研究におけるこのシステムについて把握していると同時に、秘密の知識の世界、錬金術の世界にも生きていました。アイザック・ニュートンはそのノートに、錬金術と化学錬金術に関して 120 万語も記述していました。化学錬金術については、他の何よりも科学的発見について多く発表しています。ニュートンは、ロバート・ボイルのような人々が参加する錬金術師の団体に所属していました。これは 1600 年代後半のことであり、彼らは秘密の知識の世界にいました。彼らは早まって何かを発表するのは良くないことだと信じていました。なぜなら、そのような知識は非常に危険になりかねないからです。

先に私が述べた問題は、自分の知っていることを実演する人々によって最初に解決されました。アルゴリズムを持つ数学者なら、確立された問題を解くのに、より早く計算ができる算盤の名手に挑戦するでしょう。あるいは、彼らは新しい問題を提起して、「算盤でやってみましょう」と言うでしょう。他にもオープンな挑戦がありました。評判を巡る競争は優先順位の争いを引き起こし、人々は直接的に確認をするために、他の方法を見つけたのです。

科学的不正行為に関して

次に、個人の行動レベルの科学的不正行為の問題に着目しようと思います。不正行為には、悪意のない間違いやセクシャルハラスメント、科学的不作法などは含みません。科学的不正行為には三つの形があり、それはねつ造、データの改ざん、盗用です。われわれがデータを持っている調査済みの多くのケースにおいて、盗用は中心でありません。盗用が問題になるのは、一つの領域だけです。それは、盗用が教育において認定の価値を損ねてしまう領域です。双子の兄弟の一方が賢いからという理由で、自分の代わりに試験を受けさせることができるとしたら、われわれは、本当は認定されるべきではない誰かを認定してしまうことになります。それは、教育システムが信頼の基準を生み出しているといった考えを揺るがします。

これを除けば、盗用は私の考えでは、被害者のいない犯罪です。それによって誰も傷つきません。なぜかと言うと、ほとんど知られていないことを盗む者は、マイナーなジャーナルのどこかに発表して、多分少し異なる形でそれを再出版するだけだからです。盗作者は、それを見落としていた人々に何かを気付かせるために持ってきます。なぜなら、盗用する価値のあるものは、何であれ、そのまま永遠に見向きもされないよりはいいからです。これで誰が傷つくでしょうか？元々の著者は、人々が注意を払ってくれるのですから、傷つきません。真の著者の論文は無視され、重要なジャーナルには発表できなかったというのは事実です。それがどこかに発表された結果、そのことは職務経歴書に記録できるようになったわけで、無名の存在から救われたのです。盗作者の方について言えば、その論文の価値に注目が集まれば、それを最初に手掛けたのが誰か別の人だと覚えている人が出てくるのが普通です。そうなると、盗作者はあまり長い間利益を得ることはできません。

ある種の病的なケースの場合、腕のいい盗作者は普通、自分自身の優れた成果も発表しています。彼らはこう考えます。「自分ならこれをもっとうまく書くことができるだろう。私は他の論文は全て自分自身で書いて発表しているのだから」。これは別の問題です。有名な実績ある人々の、良い実例があります。マーク・ハウザー以前に、環境医療科学およびクローン作成の分野で、ねつ造した画像を使い、クローン生物のさまざまな段階を表しているとプレゼンした者たちがいました。驚くほど多くの物理学の論文を書く、スイス・ジュネーブの IBM ラボで働く物理学者がいまし

たが、その論文全てがねつ造であることが明らかになりました。論文の結果どおりにやってみても、うまくいかなかったのです。

これについて人々は、二つあるうちの一つの方法で考えます。オープンサイエンスモデルにおいて、もしあなたが規範の一つに違反すれば、同業者から尊敬を失うという罰を与えられます。あるいは、あなたは罪を犯したことになり、罰則がやってきます。エコノミストのゲーリー・ベッカーはそのアプローチについて、「この名声を得ることで獲得する可能性のある利益は何か、どんな価値を手に入れるのか？」と言います。もし何かが注意を引いたとして、捕まらなければ、それは利益になります。一方、見つかったら何が起こるでしょうか？　あなたが苦しむ罰はどれくらいの規模になるでしょうか？　人々は、問題は検知システムが十分に強力でないことと、罰則を十分に厳しくしていないことであると述べて、これに取り組んでいます。もしそうすれば、われわれはこれが起こる頻度を減らすことができるでしょう。それが一つのアプローチです。もう一つのアプローチは、「これは間違っていることだから、やらないように教育する必要があります。それは規範違反であり、同僚が最終的に見抜くか、ラボで働く若手が気付くでしょう。彼らがもし適切に教育されていれば、これが不適切な行為だと分かるでしょうし、誰かがそれを、彼らを監督し、資金提供している人に報告すべきです」というものです。

不正行為については、入手できるデータは非常に少ないです。不正行為の申し立てがある場合、1996 年以降のデータは、現在米研究公正局となっている機関から入手可能です。彼らは最初、独自に調査しようとしていましたが、1995 年までにすぐにそれをやめました。研究を受けた機関によって実施される調査の監督だけをするようになりました。この組織は米国公衆衛生局内に設置されていたので、監視は生物医学分野の手続きや医薬品、機器に集中していました。この分野は認可申請と密接に関係しているため、多くの人が関わっています。この時期、生物医療科学の成長は著しかったため、ライバル関係も恐らく多くありました。不正行為の申し立ては数多くありましたが、キャリアを高めるチャンスもありました。生物医科学の相対的重要性は維持されています。ポストドクターやその他の研究員などの若者が流入してくるのが分かります。生物医科学分野ではポストドクターのシェアの方が高いです。研究室の規模は大きくなっています。科学の実践規模はもはや、一人の科学者やごく少人数グループの共同作業ではなくなっています。今やそれは大人数のチームで、そのチームはポストドクターに加えてその他の研究員を抱え、そのうちの何人かは文字通りの科学者ではありません。彼らは工業系や管理系の人材かもしれません。また、実証実験で患者をモニターする看護師など、研究者でも技術者でもない支援的人材もいます。

関わっている人々の地位を見るとき、最も興味深いのは、これらの人々が規範に違反しているために、オープンサイエンスモデルが行動を説明できないことです。基本的な罪と罰モデルも助けになりません。なぜなら、非常に優秀で、わずかにクレージーな人がたくさんいるからです。彼らは自分たちが見つからないと信じて、不正行為をやり続けます。それは理性的なことではありません。従って、罰を重くしても、絶対に捕まらないと思っている人々には影響しません。また、非常に初期段階にある人や、わずかにその領域に入っている人々にとっては、典型的な罰は助成金を与えないことでした。彼らのほとんどは、終身在職権はもちろん、ポストドクターとしての 2 年目すら得られないので、多かれ少なかれ研究分野から疎外されることになります。彼らは助成金取得を妨げられることが、大きな問題になっていくとは考えません。いずれにせよ助成金を得られないので、不正行為を選んでいるのです。

不正行為の罪を告発され、有罪を宣告されるのは誰でしょうか？　最も頻繁に告発されるのは若手です。大学院生やポストドクターもこのグループに入ります。全てのケースのうち、約 3 分の 2 がこのグループに入

ります。准教授や助教授、研究助手らは、境界線上にいます。正教授が不正行為で告発され、責任があると判断される可能性はかなり低いです。ここでの例外は、生物医科学における研究の産業化によって、デザイナーでなく組み立てラインで働く人々の一部に、同じ種類の疎外とアノミーの問題が発生していることです。こうした人々は指導されていません。研究室がやっていることの一翼を担っていると感じさせられてもいません。そこが重要です。オープンサイエンティストの社会へと進んでいく中において、自分の職務を適切に行わなければ、それは正当な理由を傷つけ、それがシステムの失敗になります。その原因は、ますます進む機能の専門化への対処の組織的失敗と、異なる物差しで研究を行わなければならないという点にあります。それは生物医科学分野で最も明らかであり、さまざまな方法で対処される必要があります。それは一種のパラダイムです。

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

データ駆動型の科学研究エコシステムとしてのオープンサイエンス

ー過去のコミュニティ実践事例と日本の視点

村山 泰啓

（情報通信研究機構/ICSU-World Data System）

講演要旨

今日、オープンサイエンスは、特に国際的科学政策文脈では、しばしば研究データエコサイクルのデジタル化、またデジタル基盤上での研究過程や研究アウトプットの出版・共有・配信、といった議論をされることが多い。国際的に期待されているグローバル研究データ基盤やそこで可能となる研究ワークフローを今後設計していくためには、最新の電子情報基盤のツールやノウハウが必須であるとともに、ユーザー（当初は科学者や科学業務に関わる者が主体）からみた実現可能なユースケースに基づく、より上位層のコンセプトが必要となる。過去の研究コミュニティにおける研究データ共有の実践事例は、この設計指針をあたえる上で重要である。

村山 泰啓

国立研究開発法人情報通信研究機構ソーシャルイノベーションユニット戦略的プログラムオフィス研究統括（現職）。ICSU-WDS（World Data System）国際科学委員会ex officio委員、日本学術会議連携会員、国立極地研究所南極観測審議委員会委員および重点観測専門部会長、国立国会図書館科学技術情報整備審議会委員を務める。1999－2006年は北極域アラスカにおける上層大気観測日米共同研究計画の日本側リーダーを務めた。内閣府「国際動向を踏まえたオープンサイエンスの推進に関する検討会」委員、欧州委員会「欧州オープンサイエンスクラウド高級専門委員会」委員、首都大学東京非常勤講師、公益社団法人日本地球惑星科学連合理事、京都大学生存圏研究所客員教授、などを歴任。文部科学大臣表彰科学技術賞受賞（2007年）。京都大学工学博士（1993年）。

私がこれまで関わってきたオープンサイエンスは、European Commission や、オバマ政権時代のホワイトハウスなどのデジタル・サイエンス・ベースでサイエンスのやり方を変えていこうという政策的な働き掛けと、コミュニティ側の新しいデジタル時代のオープンネスのノルムや作法、スカラリーコミュニケーションのやり方を変えていくという議論が中心でした。そこにおいて、私の議論の中心は、オープン・アクセス・ジャーナルよりは、むしろデジタルデータにありました。

自己紹介

私の研究分野は地球科学、大気科学、宇宙物理で、日本地球惑星科学連合のボードメンバーや、京都大学の客員教授を務めたことがあります（図 1）。最近では、データインフォメーションの活動に日本からも貢献したいと考えて、アメリカ地球物理学連合（AGU）や国際測地学・地球物理学連合（IUGG）といった国際学会の中で委員などをしています。日本学術会議連携会員でもあります。また、The ICSU World Data System（ICSU-WDS）の国際科学委員会 ex officio 委員で、これがデータについて取り組むきっかけになりました。他に、2016 年まで欧州委員会「欧州オープンサイエ

ンスクラウド高級専門委員会」委員を務め、内閣府「国際動向を踏まえたオープンサイエンスの推進に関する検討会」委員、G7 オープンサイエンス WG のメンバーも務めています。

科学のあり方を考える

科学に携わっていると、日本政府でもどこでも、3 年か 5 年のファンドを出すと玉手箱のようにいい成果が出て、例えば災害対策ができるというような政策決定の仕方をよく耳にするのですが、それは正しい科学のやり方なのでしょうか。科学というのは、データを集めて、論考を合理的に重ねていくというロジックの積み上げです。この煩雑な作業を通して、このあたりが確実だろうと言えるものを地道にこつこつ作り上げた結果として、今のテクノロジーがあるのです。スマートフォンにせよ、航空機にせよ、さまざまな現代のテクノロジーはそういったとても面倒な作業の上に成り立っていて、それでわれわれはベネフィットを得られています。

しかし、それを忘れて、何年間か予算を出せば科学者集団が社会に対していいものを出すだろう、金をやるから成果を持ってこいということにどうもなりがちです。不幸なことに、社会と科学のミスコミュニケーションが起こっています。それを解決するためにオープンサイエンスという考え方が役に立つのではないかと思っています。私の話はいつも抽象的なのですが、これはなぜかというと、この話を政策決定に関わる人

（図 1）

などに、将来の日本の科学のあり方を考えるための基盤にぜひしてほしいからです。

具体的なことでは、データに DOI を付ける、学協会のジャーナルのポリシーをデータに合わせて変えるといった話をすることは重要だと思いますし、そういう話を実際にしている学協会もありますが、そこだけを話しても、何のためにそれをやるのかは恐らく伝わらないと思います。私は SPARC Japan セミナーでも 2013 年から、もう 3〜4 年この話をしています。日本の中で、科学者は面倒なことをしているし、言うことを聞かないししょうがないけれども、科学に基礎を置いて政治意思決定をする世の中の中でやらざるを得ないのだということを共通理解として持てるようになればと思って、辛抱強く何度も何度もお話をしているのです。

そういった理由があるので、私はオープンサイエンスについて話すときは、ポリティカルバックグラウンドも必ず話します。その上で、サイエンティフィックリサーチのシステムは、Paul David 氏のお話と根底は同じです。そういうものの国際的なサイエンスノルムの具現化の一つとして、International Council for Science（ICSU）が開いているデータ委員会の一つが World Data System です。科学界だけではなく、社会全体で、デジタルインフラの上で社会のデータまで含めて将来像を変えていこうという動きがあるのです。こういったことを行っていく上で、もちろん具体的に学協会がどう変わらなくてはならないかということもあるのですが、実際に動きだしているということを少しずつでも議論できたらと思っています。

オープンサイエンスの国際的な動向

私がオープンサイエンスに関わりはじめたきっかけは、海外の学協会や国際会議で、研究グループ、コミュニティや社会とデータを共有していこうというオープンサイエンスの議論を聞いたとき、何だこれはと思ったことです。当時、全く聞いたことのない話だったのです。その後、G8 科学大臣会合の研究データのオ

ープン化の話が内閣府から来て（図 2)、「こんな議論があったのだけど、あなたがやっている仕事はこれに近いのではないか」と言われて話してみると、まさに同じコンセプトでした。ですから、国際社会では、アカデミーの側とポリティクスの側で、同時にオープンサイエンスが表に出てきた感があります。

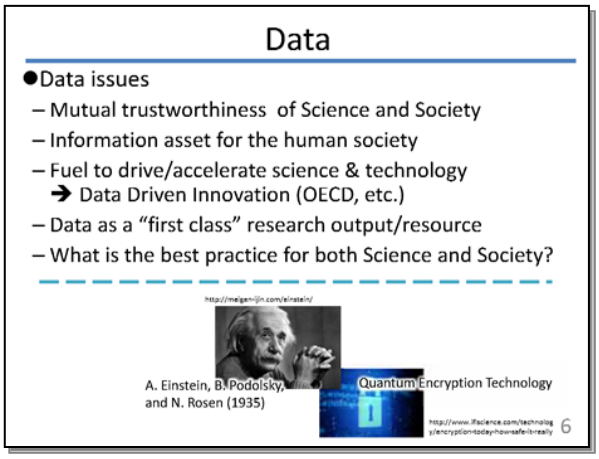
当時は、日本でオープンサイエンスの話をしたら、「何でこんな話をするのだ」と言われましたし、いろいろな研究機関の方にお話をしても、「何それ」と引かれるのが関の山だったのですが、あれよあれよという間に重要なトピックになってきたのではないかと思います。

2016 年の G7 Science Ministerial Meeting では、林和弘さんと一緒に話題提供講演をさせていただきました。オープンサイエンスのワーキンググループでは、G7 の合意の下に、インセンティブやデータインフラストラクチャーの重要性についての議論が進んでいます。昨年はイタリアで G7 があり、その次に向けてどうしていくかが考えられています。

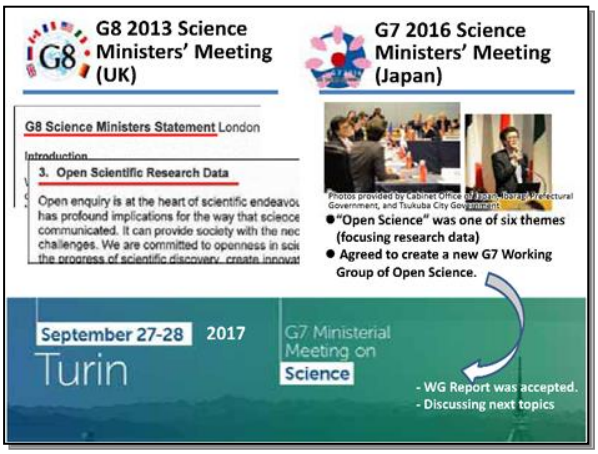
でも、何か偉い人たちが集まる場でこういうことを言っているから、オープンサイエンスはきっといいことがあるに違いないと思っていても仕方がないのです。研究者や図書館員、出版社、学生、研究機関、企業など、いろいろな人にとってオープンサイエンスは重要だけれど、トップダウンで進めなくてはいけないことも少しはあって、そこがないとどうしてもこれは進まないなという部分をきちんとトップが集まる場で決めなくてはいけません。そうでないことをむやみにトップが集まる場に持ち込んでも意味がなく、お祭り騒ぎになるだけであって、本当に進みはしないので、そこのところは重要だと思っています。国内で議論していると、そこをあまり理解いただけていない場合があるのですが、何でも持ち込めばいいという話ではないのです。

図 3 は、2016 年の G7 Science Ministerial Meeting で出したスライドの一つです。これは単に政治家にメッセージを伝えるためだけのものです。データは社会と科学の相互信頼の問題であって、データはテクノロジーや科学を進めるための fuel（燃料）で、データは研究の成果としても非常に重要な生産物なのだということを改めてお伝えしました。

日本では内閣府が、「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」を開催し、オープンサ



（図 3）



（図 2）

（図 4）

イエンスのナショナルプリンシプルのようなものをレポート「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」にまとめました（図 4）。ウェブに英語版も日本語版も出ています。これは原山優子議員が Council for Science, Technology and Innovation（CSTI）の代表として作り上げたものです。座長は九州大学の元総長の有川節夫先生で、非常に重要な議論をされたと思います。私もこの委員会に入らせてもらって、作業に関わりました。

知識の共有と再現性

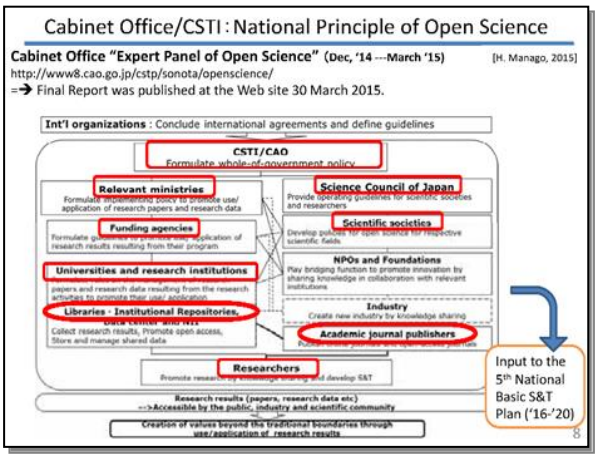
本日は出版社やライブラリアンの方が多いと聞いたので、そこの論点も少し入れていこうと思っています。図 5 はその内閣府のレポートの中に出ている図です。当然ながら、内閣府があって、その関連省庁があって、ファンディングエージェンシーがあって、研究機関があるのですが、図書館やリポジトリもきちんと位置付けられています。これは本当に重要なのです。同時に、サイエンスカウンシル、学協会、ジャーナルパブリッシャーが、サイエンスのシステムとして入っています。

国立大学が比較的メジャーな日本で、ここで省庁から国立大学に話が降りてきて、法的な強制力でもってデータをオープンにするということになると、サイエンティストの振る舞いでは全くなくなります。この、カルチャーを変えていくという部分を抜きにして、強制力が働くのは望ましくない、だからこそ、ここには法的な強制力でないサイエンスソサエティが入ってこなければいけないというのが私の意見です。

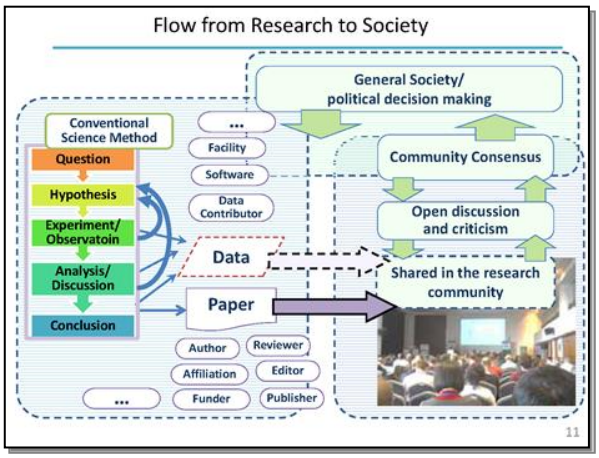
伝統的なサイエンスのメソドロジーは、クエスチョンからハイポセシスが出て、分析してコンクルージョンを得るというもので、これが学生向けのテキストなどに載っています（図 6）。ここから出てきた結論を学協会に発表することが、改めて知識を共有する、サイエンスのオープンネスであり、ディスインタレステッドネス（公平）です。オープンなディスカッションと相互批判をすることで、みんなの共通理解ができていくというプロセスを経ることが重要だということです。そのため、企業の中で閉じて研究するような場合は、みんなの共通理解になりません。それをオープンにして進歩をさらに進めていく上で、論文は非常に重要な役割を果たしてきたのですが、ここで今さらながら、現代の論文というのは検証可能かという話に立ち戻ります。

論文に図を載せて、過去 20 年でこのような気候変動があった、だから次はこの時期が問題だと言うけれど、その図は本当に正しいのか、著者が徹夜して書いたものではないということがきちんとデータでバックアップされなければならないということです。

ここでデータが非常に重要なアクターとなり、データサイエンスという言葉も出てきます。私の感触だと、私が博士学位を取ったのは 1993 年ですが、当時既にデータにものをいわせて博士論文を書いているのです。当時、既にそこではデータサイエンスは成立していたのではないかというのが私の意見です。それ以前の私



（図 5）



（図 6）

の先生が論文で使っていたデータは、論文中にデータテーブルで書ける程度の量だったのですが、私ぐらいの世代からは、もうそのような量のデータではありませんでした。統計処理して抽象化された図を出すしかなかったのです。そこでモードチェンジが起こっていたのではないかという気がします。

本当はいろいろな付帯物が一緒に出なければ結果は再現できず、データで全て終わりというわけでは全くないのですが、今はデータを最初の第一歩として、データだけにまずは集中してデータの再現性がいわれているという状況です。

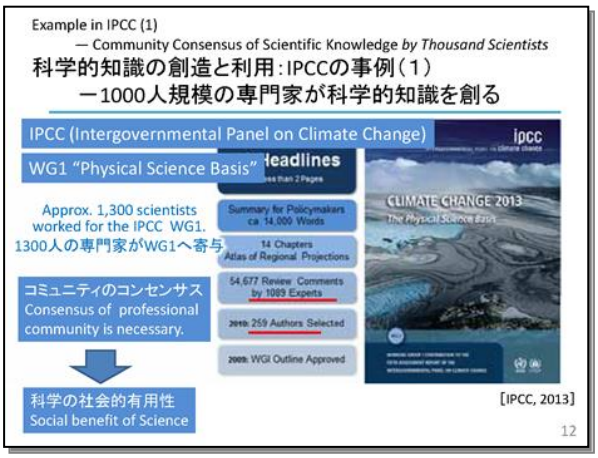
ソーシャルシステムの中の科学

Intergovernmental Panel on Climate Change（IPCC）のレポートは社会に重大なインパクトを与えてきたということは何度もいろいろなところで話しているとおりです（図 7）。1,000 人規模の専門家が集まって一つの報告書を書いています。このようなコミュニティコンセンサスがつくれてこそ、科学的アウトプットは社会にとって重要な知見であり得るし、また、それで実際に政策決定者が影響を受けて社会が変わっていき、社会のベネフィットが端的に表れると言えると思います。

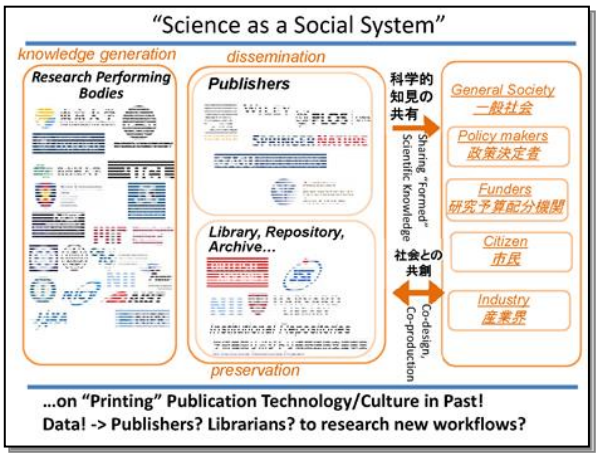
研究活動を行い、それを発表し、蓄積し、引用したり再利用したりして、エコサイクルを回すことで、サイエンスの営みが成立してきたのですが、そこでつくられた科学的理解のコンセンサスが社会に提供されて初めてその有用性が分かります（図 8）。これは、貴族の時代のように、「この知識はすごいだろう」という話で済むのであれば要りません。いい知識さえ出てくれば自慢できるというものではなく、一般社会へ伝わらなければ、科学的知見が次の世代に伝わらないし、次の世代で教育されていかないと次の研究が進みません。われわれの科学というのは、研究機関から、出版社・図書館・リポジトリなどを経て、一般社会で使われて初めて、ホビーではなく、社会の営みになっていくのです。

それをする上で、われわれが研究して、出版社・図書館・リポジトリが付いてきてくれればオッケーかというとそうではありません。われわれが研究してパブリックファンディングを受けられるように、きちんと一般社会に情報が渡るためには、研究機関と、重要な科学のパートナーである出版社・図書館・リポジトリが共に働ける状況でなければなりません。そういうシステムをわれわれは意識せずにつくってきました。明治に近代科学が日本に輸入されたときに、法律で大学図書館を必ず設置することになったのも、まさに近代科学の方法論を導入した現れではなかったかと思います。それが今、PDF で論文が取れるからといって、図書館が大学から消えていいのではないかという議論になるのは、サイエンスのオープンネスやフェアネスなど、原理に基づいた思考がアジアの国でなかなかされていないからではないかと思います。

出版社・図書館・リポジトリは、印刷技術の基盤において育ってきたカルチャーでできているのですが、



（図 7）



（図 8）

それがデジタルになった瞬間に崩壊したのでは、近代科学の考えそのものが崩壊しますから、それでは大学も研究機関も 10 年、20 年たったら用無しになってしまいかねません。科学を守ることは、われわれ自身を守ることでもあるし、国を守ることでもあるのですが、同時に世界の文化を守ることでもある、そういった意味で、われわれが日本の中で近代科学のやり方をどうやって存続させていくかということが、これから非常に問われる問題ではないかと思います。

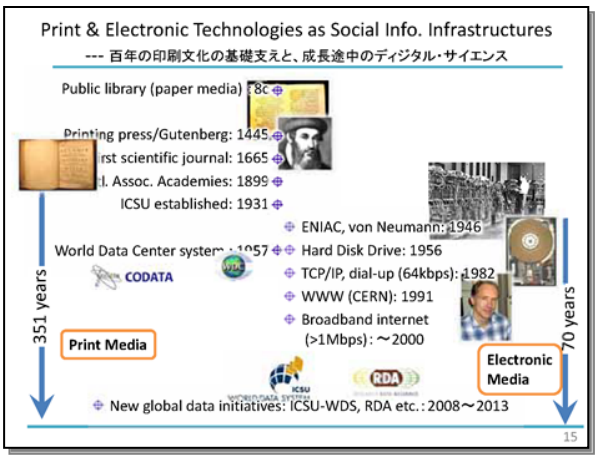
印刷文化は、300 年以上という非常に長い間、科学を進めるための手段として役に立ってきましたが、デジタルメディアというのは非常に短い 100 年未満の歴史の中でここまで普及してきたし、みんなが関心を持って使っています（図 9）。しかし、ソーシャルシステムとして、プリンティングテクノロジーと同じぐらい信頼できるのかどうかというのはこれから問われる問題で、そこをわれわれが今解決しなければいけません。研究者はデジタルメディアを使って、テクノロジーを今までつくってきたのですが、それを社会基盤にしていくところで、研究者だけではもう立ちいかなくなったのです。だからこそ Research Data Alliance のように、プラクティショナー、ライブラリアン、いろんな方が入ってきて、ポリシーメーカーも入ってきて、一緒にソーシャルインフラストラクチャーにしていくためにどうしたらいいか、国際的にみんなが悩んでいます。その中で、アカデミーからも World Data System など国際組織もいろいろ入って議論しているのです。

WDS の歴史は前史からいくと 60 年ぐらいあるのですが、今のデジタルインフラストラクチャーに対応するためにいろいろなことを変えなければいけないし、元々地球科学が中心だったのですが、メディカルからエンジニアリング、ソーシャルサイエンスまでいろいろなものが今入ってきて、データ基盤をどうするか考えているところです。

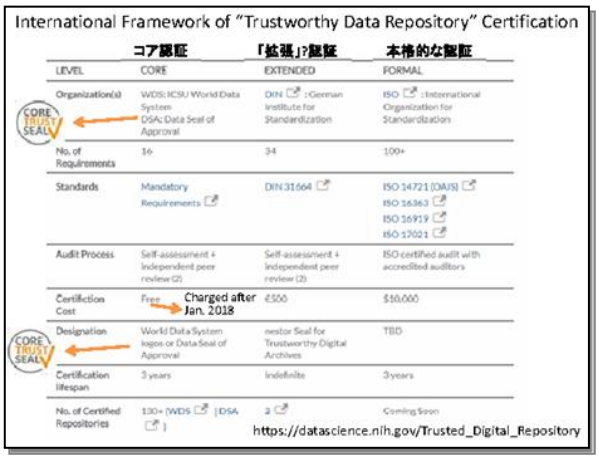
データを蓄積するために、CoreTrustSeal という新しいデータリポジトリ認証の仕組みを WDS とヨーロッパの Data Seal of Approval という機関で立ち上げました。要するに、サイバーセキュリティのような信頼性と別に、データそのものの信頼性を長期間担保するために、国際枠組みが必要だと考えて、そのための取り組みを行ってきたのです。

図 10 はアメリカ国立衛生研究所（NIH）のホームページに載っていた図で、ISO と WDS はそれぞれレベルの違うサーティフィケーションなので、国際認証がなくてはいけないという話をしているということです。

そういったことを具体的に社会の中でインプリメントするために、European Commission は digital single market strategy、日本政府は Society 5.0 を考えているのですが、やるとしたら 10 年、20 年、30 年かかり、カルチャーもインダストリーもエデュケーションもみんな変わっていかないといけないのだろうから、5 年間の five year strategic plan で完成する話ではないと思います（図 11）。それはインターネットの世界でも当然



（図 9）

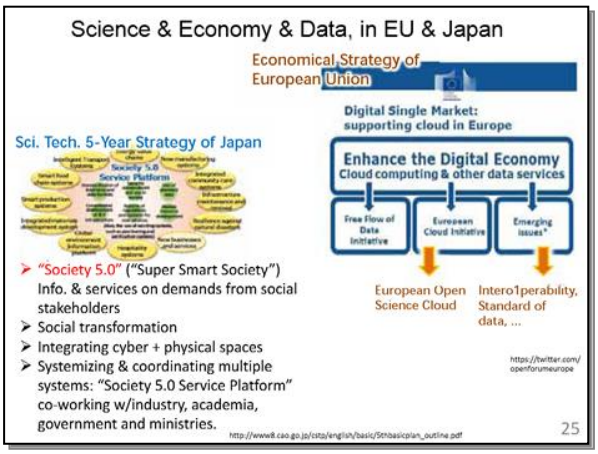


（図 10）

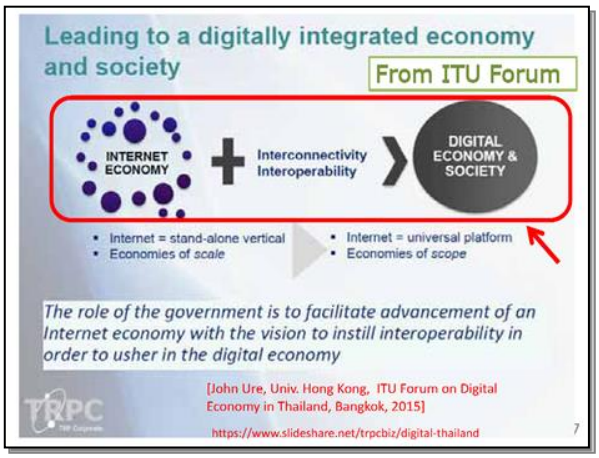
議論されています。

インターネットエコノミー＋インターコネクティビティ、インターオペラビリティで、デジタルエコノミーになるというのは、ITU Forum、情報通信の世界でも議論されています（図 12)。

研究者は、ファンディングや省庁を通じて政治的な合意の影響を受けますが、そこに通底してインターナショナルな共通理解がなければ、サイエンスとしては機能しません。それを機能させていくために、リサーチ・アウトプット・マネジメントという形で、図書館・出版社・アーカイブ・リポジトリといったものが、単なる研究のサポーターではなく、社会との関わり全てを支える骨組みとして、研究者と共に生きていってほしいと思います。これらが不可欠なパートナーとして存在し続けていくことが、科学そのもの、現代文明そのものの基盤になるのだろうという枠組みを描きながら、日本の将来が議論できるといいのではないかと考えています。



（図 11）



（図 12）

第3回 SPARC Japan セミナー2017「オープンサイエンスを超えて」

質疑応答・ショートディスカッション

林　和弘　（科学技術・学術政策研究所）
Paul A. David（Stanford University）
村山　泰啓　（情報通信研究機構/ICSU-World Data System）

●**林**　10 分ぐらい質問とディスカッションの時間を取りたいと思います。正直、哲学的な話もあって、難しい内容もあったかと思います。そのクラリフィケーションを含めてまずご質問をお受けしたいと思うのですが、いかがでしょうか。

●**フロア 1**　Professor David にお伺いします。冒頭で、オープンサイエンスというのは、従来型のマーケットバリューをインセンティブとしていないことから、ペキュリアーなシステムであるというお話がありました。その後、オープンサイエンスにおけるインセンティブというのは、レピュテーションに基づくものであるという示唆もありました。

われわれはマーケットバリューに基づくリワードシステムとレピュテーションシステムを対立概念と捉えてオープンサイエンスを推進していくべきなのか。もしくはレピュテーションというのは最終的にはマーケットバリューに類するようなリワードシステムにつながる途中段階だと捉えてシステムを構築していくべきなのか、お考えをお聞きしたいと思います。

●**David**　それは多くの混乱や、よろしくない政策を招く問題ですね。EU などで広がっているのは、学術機関内にさまざまなインセンティブを混在させるやり方です。公的資金、あるいは民間資金の場合もありますが、その資金を使って、そしてその機関の研究人材を活用して得られた研究結果に関する特許出願を認めるのです。

そうなるとわれわれは、その組織内においてはどんな研究が好ましいのかを決定しなければなりません。どのような研究に、新しい建物や施設が与えられるべきなのか。どのような研究に、新しいポジションや、より多くの大学院生を配置するべきなのか。院生は、科学系の学部において労働力の一端を担っていますからね。そしてこうした需要は多くの場合、収益性があるかどうかの認識によって変わっていきます。というのも、役人や政策立案者は、産業界の声に耳を傾けた上で、「産業界が、これが私たちに必要な研究だと言っている。これこそが、科学の進歩が見られる方向性であり、そしてそれが国民に大きな利益をもたらすだろう。なぜなら、それが産業界に大きな利益をもたらすからだ」と言ってくるからです。

例えば、軍産複合体に関する有名なエピソードがあります。ゼネラル・モーターズ（GM）のトップがこういう発言をしたのです。「GM にとって良いことは、国にとっても良いことである。GM にとって良くないことは、国を損なうことである。よって当社は、収益性と当社のシェアホルダーの将来的な富を追求することによって、国益を追求する」と。しかし、その実、シェアホルダーというのは国民のごく一部になってきており、彼らは必ずしも科学が進歩する可能性を認識しているわけではありません。そして、その進歩した科学が今後、彼らの産業に取って代わる可能性もあるわけです。しかし彼らは、自分たちの産業を維持したいと考えています。組織が大きくなればなるほど、内部組織は内部からの組織の管理よりも、環境の管理に

より依存するようになります。外部からの情報を処理するための構造をつくり、そして情報を処理するために、各種のコードを課す必要が出てくるわけです。そのコードには二つの働きがあります。コードは使えば使うほど良くなります。物事をより分類しやすくなり、その方法も洗練されていくのです。しかし、いつしか分類に適合しないものは、目に入らなくなっていきます。そしてそれは大抵、新しいものなのです。新しいものには、しばしば危険が潜んでいます。外側から自分たちの分野に入り込んでくるもの。姿を現しはじめているけれども、どこか別のところで追求されているので、それに気が付くすべがなく、知らなかったものなどです。こうして比較的大きな組織は、物事についての情報を得ようとするのではなく、自分たちの立場を守るための戦略を取るようになっています。政治的手段を使ったり、公的資金をコントロールしたりして、既存の構造を守ることに時間を費やします。つまり、守りに入るわけです。組織が大きく、政治的影響力が強ければ強いだけ、その組織のリソースは、将来の可能性に対応するためではなく、過去の遺産を守るために、より多く割り当てられます。知識を理解せず、リソースを制御することによって、レガシー的な立場を守っているのです。

バイドール法の問題点については、EU や、他の国々でも広く誤解されていました。フランスで政府関係者と話をしたときも、彼女はこの法律にどのような成り立ちがあったのか、全く分かっていなかったのです。しかしながら彼女は「*Le nouveau défi américain*」すなわち「新しいアメリカの挑戦」だという言葉を返してきました。彼らは他の国々を追い抜こうと何かしており、フランスにもこのシステムが必要だということでした。しかしながら、アメリカにはシステムがありません。アメリカ政府とその各部門は混沌としており、互いに対立し合うこともよくあります。しかし、アメリカには莫大なリソースがあるのです。ですから、この混沌としたシステムの中で、多くを無駄にすることができるのです。これは、リソースに恵まれていない国には、まねのできないことでしょう。何かがうまくいくらしいということになると、他の人たちがそれをまねしようとします。これは、政策を推進する方法としては、あまり良いものとは言えないでしょう。

彼らが間違っている点の一つが、大学というシステムの中に、相反する動機を投入してしまったことです。つまり、これは市場志向型の研究活動という方向性で、サイエンスのコミュニティの中に、研究の方向性として、二つの異なる報酬制度が存在することになってしまいました。特許を出願するか、あるいは主要なジャーナルに論文が掲載されるかという違いです。後者の場合はピアレビューがあり、論文が発表されるまでには長い時間がかかります。しかし特許の場合は、行きたければすぐに特許庁に行くことができます。しかし、こうした二つの報酬制度を共存させるというやり方は、どちらのシステムのパフォーマンスも低下させる結果に終わる傾向があるのです。人々は研究にもっと時間を費やせるようにと、教育現場から脱出しようとするようになります。というのも、たくさんの特許の申請を大急ぎで行わなければならないからです。前に申請した特許の立場を守るための特許を急いで出さなければならない。つまり、ロジックが変わってしまうのです。そしてその結果、学部内で争いが起こったり、リソースの取り合いが起こったりする場合もあります。また、大学の教育的機能から、リソースが奪われる傾向もあります。自分たちが合格させたいような学生を形成し、研究のサポートに役立つ学生が好まれるようになるのです。そして彼らをたくさん合格させたら、彼らに研究の場を用意し、面倒を見てやらなければならないのです。

ですから、システムがすべきことは、動機が異なるサブシステム同士を交わらせることです。そうして互いに取り引きをさせ、もう一方のサブシステムのためにできることもやらせましょう。そして、ビジネスの世界が、公的機関に資金を提供できると政府に伝えているようなことをやって、間接的に資金提供を受けるのです。この問題についての私のモットーは、二つの

相反するシステムがある場合、その二つが密接に結び付いていなければ、互いに協力し合えるというものです。しかし、交わらせるのはいいですが、混ぜて一つにしてはなりません。混ぜて一つにするというのは、一つの組織に異なる二つの報酬制度を導入するということです。産業界の人たちは「人を募集するためのシステムと、集めた人材を成長させるためのシステムの二つがあると、結局は機能不全に陥ってしまう」と言うでしょう。しかし、必ずしも人材を集めたときに使ったのと同じ基準で、彼らを成長させる必要はないわけで、その逆も同じです。自分たちが成長させたいと思うような人材が欲しいというような観点で、人材を募集することなどないのです。次の商業戦争の際には利用して、製品が完成したらお払い箱にするような、使い捨て要員を採用するのではいけません。

ご質問では、これまで物事がうまくいかなかった部分の、重要な問題に触れてくださいましたね。物事をうまく運ぶには、公共部門と民間部門のメリット、そして経済学者が提供できるであろう枠組みを理解する必要があります。

●林　最近のはやりだと、altmetrics などでよく行われる議論に関しての解説を一気にしていただいたような気がします。インパクトアセスメントというのはさまざまなベクトルがあって、それをどう数えるかというのは、どういうリワードシステムと関連付けるかという話になるのだと思っています。

●フロア 2　John Wiley & Sons の者です。David 教授、村山教授、ためになるプレゼンテーションをありがとうございました。私たちは出版社ですので、村山教授の社会システムとしてのサイエンスについてのプレゼンテーションで触れられていたことについてコメントしたいと思います。教授のプレゼンテーションでは、オープンサイエンスを推進していく中で重要な役割を果たす、全てのステークホルダーが示されていたと感じました。出版社としての観点には、二つの角度があると思います。一つは、助けるということです。私たちが取り組んでいることには、二つの課題があります。第一の課題は、行動変化を起こすことです。これは出版社もそうですが、主には研究者の行動やインセンティブの部分を助けるということになると思います。そして第二の課題は、インフラや持続可能性に関わるところにあります。

第一の課題については、確かセオドア・ルーズベルトがよく「アメとムチ」アプローチということを言っていたと思います。オープンアクセス（というよりはむしろオープンサイエンス）への移行において、論文や出版といった、アカデミアや進歩の重要な部分は残っていく、そうした未来の一環としてオープンデータを求めることはできるのではないでしょうか。ワイリー社も、私の知っている他の出版社も、そういったことを試しています。数百ものジャーナルを網羅する、オープンデータ化の義務付けに向けて準備しているのです。しかしそうなった場合に、われわれが懸念しているのが、忙しい研究者の方々は抵抗されるのではないという点です。そうなっても、当社のジャーナルに論文を掲載したいと思ってもらえるだろうか？　そこで分かったのが、投稿プロセスの簡略化が重要だということです。手間の掛からない投稿プロセスの一環として、ドロップするだけで、適切なリポジトリに送られるといったような仕組みが重要になってくると思われます。

第二の課題はインフラ周りです。私たちのシステムを、オープンかつ連合政府のインフラと相互運用可能なものにいかにしていくのか？　例えば世界の地球科学者を例にとると、彼らが特定のリポジトリを好んでいるとするならば、どうすればそれを一つにリンクされたアプローチに結び付けることができるのか？　これは全てのステークホルダーにとっての次なるステップの一部であり、そして出版社が手助けできる部分であると思います。

●村山　ありがとうございます。私も同意見です。プレゼンテーションでは、カルチャーの変化や、研究者および組織の行動を促すための具体的な手法には触れませんでしたが、おっしゃるとおりだと思います。確かに「アメとムチ」には効果があります。ここで私が一番懸念しているのは、日本のコミュニティです。日本では、政策をパッケージとして設置することが簡単ではありません。学会によっては、データリポジトリに関しては懸念しているけれども、ジャーナルの掲載方針は気にしていないというようなところもあります。そこで必要なのは、掲載方針やデータ、およびデータの保管、そして相互運用可能なデータシステムといった、エコシステム全体で協調していくことだと考えます。そうすれば、研究機関がデータをリポジトリにどのように格納し、いかに他の機関とやりとりしていくかといった点も変わっていくでしょう。それはつまり、相互運用可能なインフラストラクチャーの問題です。従って解決不可能な問題ではありません。これまで私たちのデータはオープンになっておらず、そして論文もオープンになっていませんでした。論文を読もうと思ったら、お金を払って購入する必要がありますし、論文によっては、例えば NASA の古い図書館や、米軍の図書館に保管されているようなものもあります。しかし、そうした状態においても、サイエンスは進歩してきました。そして私の個人的な見解としては、研究者に全てをオープンにさせる必要はないと思っています。情報の流通を良くして、それを互いに共有するための適切な方法は、自分たちで選択することができるのです。そんな中での出版社のご尽力は非常にありがたいことです。

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

オープンリサーチを可能にするには

Heather Joseph
(SPARC North America)

講演要旨

世界的に資金提供者は研究プロセスの多くの面での情報公開により重きを置くようになっている。それは文献やデータセットのオープンアクセスを要求することから、プレプリントの活用の促進、オープンピアレビューを活性化するといったことまで、そのようなオープン化を可能にする資金提供が挙げられる。このようなオープン化の活動を支援する政策の採用増により、研究事業を行うあらゆるステークホルダー、つまり個々の研究者から研究機関までが全体として、新しい複雑さを多く抱えることになっている。本講演ではさらに多くのオープンリサーチ事業に向かっている動きが示す挑戦と可能性を探り、移行プロセスを促進しスムーズにするための戦略を提起する。

Heather Joseph

2005年からSPARCエグゼクティブディレクター。就任以来、電子文献、データ、教育資源のオープンシェアリングための新しいモデルづくりを支援することにSPARC事業の重心をおいてきた。その指導力を発揮し、SPARCは効果的なオープンアクセスポリシーと実践を啓発する国際的指導力をもつ団体として広く知られるようになった。ワシントンDCを拠点として、オープン化政策に関する課題に取り組むアメリカ合衆国の政策立案者らのアドバイザーとして日ごろから活動している。Commerce Data Advisory Councilのメンバーとしては、オープンデータに関する政策について米商務長官に助言する立場にある。また国立衛生研究所、2016年オープンデータに関する大統領移行チーム、米国科学アカデミーに対しても同様の役割を担っている。

主催者の皆さま、この度はこのセミナーにお招きいただきありがとうございます。このセミナーに参加できて光栄ですし、皆さまと一緒に、特に SPARC Japan の同僚たちと一緒に時間を過ごすことができ、うれしく思っております。私は、次の 3 点、オープンアクセスへの動き、進歩、そしてオープンアクセスの採用にまつわる課題を、オープンサイエンス全体が直面する課題を考察する一つの方法として取り上げ、お話ししたいと思います。

私たちは、論文へのオープンアクセスはオープンサイエンスの要だと考えています。オープンサイエンスのエコシステム全体においてはほんの一部分にすぎませんが、非常に重要なパズルの 1 ピースです。SPARC は学術・研究図書館の国際的な提携で、研究と教育の世界でオープンをデフォルトとすることを使命としています。「オープンをデフォルトにする」というのは、やむを得ない事情がない限り、研究結果や授業で使う資料を共有するということです。これは基本的に、現状とは正反対です。

北米の SPARC では、学者や科学者が研究を行うために最も猛烈に必要とする資料へのアクセスを開放することで、オープンな研究エコシステムを支援することに第一の焦点を当てています。これには、論文、研究データ、教科書などの教材、コースウェアのシラバス、学習目標が含まれます。

ワールド・ワイド・ウェブの可能性

私たちは、2018 年には、ワールド・ワイド・ウェブという素晴らしいプラットフォームと指先に広がるインターネットのおかげで、効果的な科学、研究、学問活動を自動で行えるようにするために必要な全てのものに、自由にアクセスできるようになると考えています。デジタルネットワーク技術の登場により、私たちははるかに多くの情報にアクセスできるようになり、その情報を使ってはるかに多くのことをするチャンスが得られるようになりました。これは私たちにとって簡単であるはずなのですが、理論上の話にとどまっています。インターネットがもたらす希望にもかかわらず、優れた研究や優れた科学活動を行うために自由に使用できることが最も猛烈に必要とされる資料には、制約のある方針やアクセスへの障壁、価格の障壁、デジタル環境で情報を利用・再利用する上での障壁が重なり合っています。

アクセスへの金銭的障壁

SPARC は図書館の会員組織なので、現状を変えたい、変えなければならないと考えはじめた一番の理由は、金銭的障壁、キャンパス内の同僚に情報へのアクセスを与えるコストに対応するためでした。こういった金銭的障壁が一番大きいのはジャーナルです。図 1 の表は、2016 年のジャーナルの平均価格を米ドル建てで示したものです。基本的に、特定の学問領域のジャーナルの記事 1 年分のアクセスを獲得するのにどれほどお金がかかるか、お分かりいただけると思います。平均は 1,788 米ドルで、一連の論文への 1 年分のアクセスの料金としては大変な額です。

個々のタイトルのレベルでは大きくなくても、積み重なれば巨大なビジネスになります。これらのジャーナルから得られる利益は年間およそ 100 億ドルに達します。これはもうけになる産業です。従って、このような資料にアクセスしたいと望む人、そしてこのような資料へのアクセスを獲得する任務を負っているライブラリアンにとっても、大変な障壁、金銭的負担となっています。

コストの壁は、単なる理論上の壁ではありません。多くの研究者が、「高いけど私たちが必要なもの全てへのアクセスは図書館が手に入れてくれる。私たちには、必要なアクセスにお金を出してそれを手に入れる能力がある」と考えています。確かに、大学の教職員が、重要な論文に十分アクセスできないと感じて大学の建物の前で抗議をするという姿を私たちが目にすることはありません。しかし、彼らは確かにアクセスが不足していると感じており、私たちも毎日それを目の当たりにしています。

今ここにいる皆さんは全員こういう状況に直面したことがあると思います。皆さんは、自分が関心を持つテーマについて研究しています。自分の研究に役立つ情報を得るために、使いたい論文をグーグルで検索します（図 2)。私の息子は 1 型糖尿病を患っているので、私は息子の血糖値が夜の間に低下しないようにす

ISI Indexes 2016 Cost By Discipline

Discipline	Average Cost Per Title
Chemistry	$5,105
Physics	$4,508
Engineering	$3,244
Biology	$3,104
Food Science	$2,729
Astronomy	$2,718
Botany	$2,418
Geology	$2,400
Technology	$2,239
Zoology	$2,221
Math & Computer Science	$1,895
Health Sciences	$1,801
General Science	$1,717
Geography	$1,713
Agriculture	$1,687
Business & Economics	$1,474

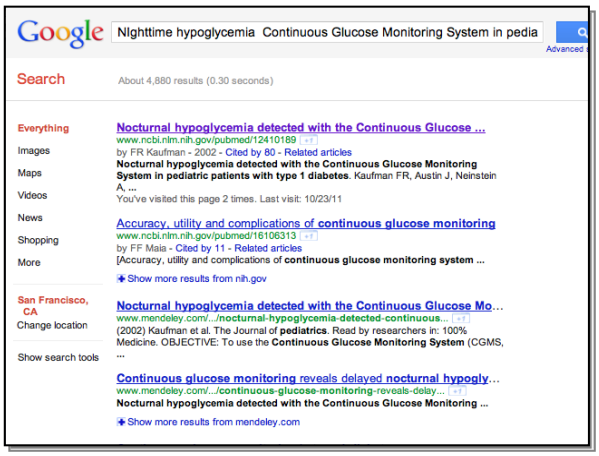
Discipline	Average Cost Per Title
Military & Naval Science	$1,063
Psychology	$1,020
Sociology	$1,004
Education	$978
Social Sciences	$907
Political Science	$820
Library Science	$774
Recreation	$747
Anthropology	$513
Law	$475
History	$434
Philosophy & Religion	$433
Arts & Architecture	$432
Language & Literature	$379
Music	$293
General Works	$263

Average: $1,788

Source: Library Journal 2016 Periodicals Pricing Survey

"Tracking the Ecosystem | Periodicals Price Survey 2016," by Stephen Bosch and Kittie Henderson, Library Journal, April 21, 2016.

（図 1）



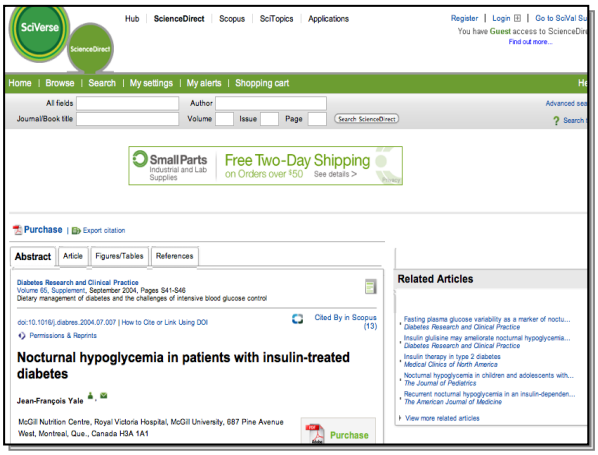
（図 2）

る方法を求めて、いつも最新の情報を探しています。自分の関心を引きそうな論文が見つかります（図 3)。要旨を読んで、良さそうだなと思います。読んでみようとしますが、論文の全文にアクセスしようとすると、ペイウォールに突き当たります（図 4)。私たちは全員、そんな経験をしています。

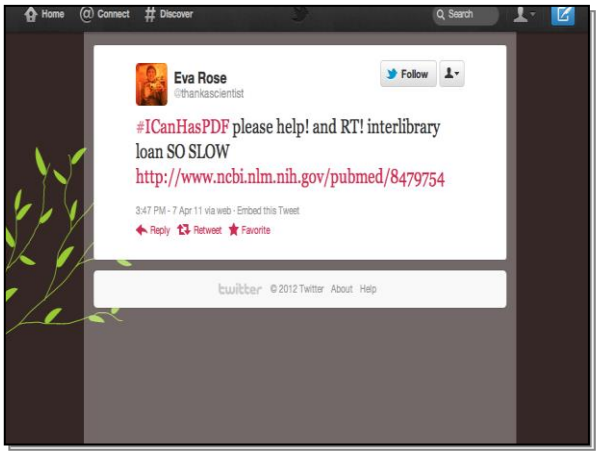
アクセスへの金銭的障壁の克服

このような事実にもかかわらず、私たちは抗議していません。私たちがしていることは、オープンな世界に向かって変えるべき行動を理解するためにとても重要です。ライブラリアンとしては、研究者が必要とするコンテンツにアクセスできない場合、私たちに ILL（図書館間相互貸借）で論文のコピーを入手してほしいと頼んできてくれればいいのにと思います。そういう場合もありますが、大抵の場合は直接著者に掛け合うか、購読している他の組織の同僚に頼んでコピーを入手します。若い学生やキャリアの浅い同僚の多くは、#ICanHasPDF のハッシュタグを使って Twitter 上でコピーを探しています。これはとてもクリエイティブな方法です（図 5)。

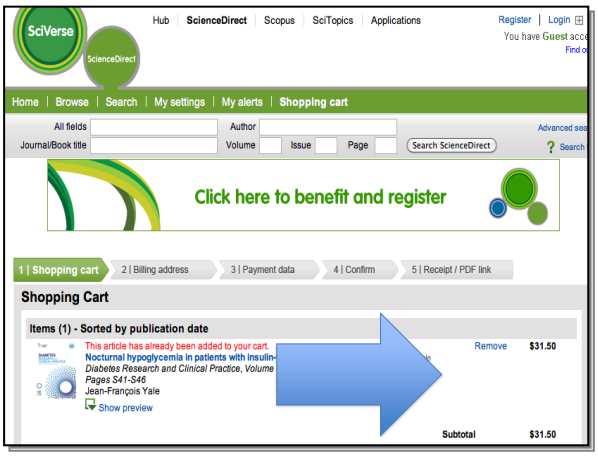
リポジトリやウェブを検索して、ペイウォールの向こうにある論文の無料のコピーを探すちょっとしたアプリを作る人も出てきています（図 6)。巨大な海賊版論文の闇市も生まれています。Sci-Hub は、学術ジャーナルの 5,800 万本の海賊版論文を集めた完全に違法なデータベースです。Science Magazine が、誰が違法論文にアクセスしているのかという調査を行いました（図 7)。ご覧のとおり、赤くなっているところからアクセスがあったのですが、世界中に散らばっています。これは水面下の問題です。キャンパスで抗議をすることはなくても、毎日の生活に確かに影響が見ら



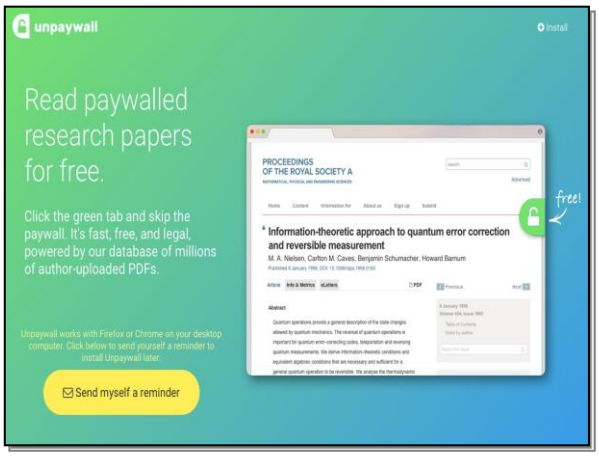
（図 3）



（図 5）



（図 4）



（図 6）

れています。

もっとまずいことに、私たちは科学者、学者、研究者から、ペイウォールに突き当たると、自動的に頭の中で計算をして、その論文の無料のコピーを入手しようという努力はせずに、その論文を飛ばしてアクセスできる論文に移ってしまうということを繰り返し聞かされています。このことは、本当に知る必要があることよりも、アクセスできるものをベースに研究が行われているということを意味します。さらにまずいことに、学生に教えたり他の人たちに情報を提供したりするときにも、本当に知る必要があることではなくアクセスできるものをベースにしてしまっています。

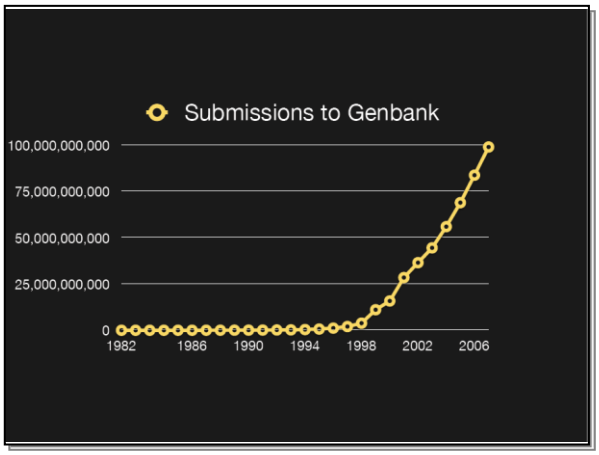
限定的な実用性で悪化するコストの問題

基本的には、私たちはインターネットが可能にしてくれたこの素晴らしい世界において、素晴らしい科学活動や研究を行うために必要な基本的な情報を入手するため、その場しのぎの対応をすることを余儀なくさせるシステムの中で活動してきました。考えられるソリューションとして、オープン化を考える際には、これらのデジタルの資料へのアクセスを入手することだけが重要なのではないということを認識すべきです。私たちはデジタル論文とそれを支えるデータにアクセスできるようになり、デジタル環境で完全に、また自由に利用できるようになることを求めています。

一例として、ある学問領域がデジタルでの研究の実施に移行すると（ヒトゲノムのデジタル化によって遺伝学やゲノミクスで起こったように）、付随するデータや情報の量がそれに合わせて急激に増加します（図 8）。ゲノムが 1990 年代後半にデジタル化されると、ゲノムのデータベース Genbank に登録されるデータの量が急増しました。これはゲノミクスに特有の現象ではありません。デジタルの動作環境に移行した学問領域では全て、このような情報の増加を経験しました。そして経験し続けています。

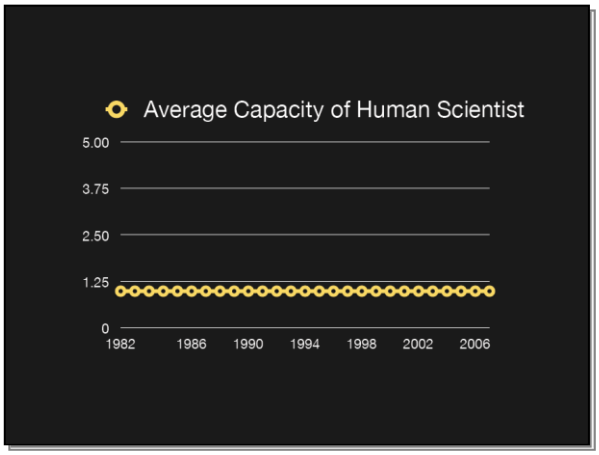
個人として、また個々の科学者としてこのような情報を処理する私たちの能力は変わっていないことを覚えておくことが非常に重要です（図 9）。私の友人、キャメロン・ネイロンは、以前よりはるかに多くの情報を処理することが求められているが、自分には一つしか脳がないと言っていました。

このことは、私たちがコンピュータを、科学・学術論文、そしてこれらの論文の土台となる付随するデー



（図 8）

（図 7）



（図 9）

タの読者のカテゴリーにする必要があるということを意味しています。私たちの目の前にある、増加する情報の意味を理解するためには、そうしなければなりません。私たちは、デジタル資料が持つ情報の力を本当に解き放つため、論文のサブセットやアクセスできる論文だけでなく、研究のための情報を得るのに必要な全ての論文や付随物に対して、テキストマイニングやデータマイニング、コンピュータ分析ができるようになることを求めています。

また、私たちはますます政治的な環境の中で活動していて、科学の有効性や価値が問題にされています。私たちは結論や論文の有効性を実証できるようになりたいと思っています。科学の有効性と真実性を定期的に確認し、再現し、承認するには、データとしての論文、そして根拠となるデータへのアクセスが必要です。これは本当に重要なことです。

私たちのコミュニティをオープンアクセスかつオープンな研究環境に移行させる動機として、私たちがSPARC でいつも話していることは、研究結果と付随する教材を共有するシステムを最適化して、科学活動をしている、そして科学を利用している人たちのニーズにもっとうまく対応できるようになるということです。このシンポジウムではオープンサイエンスについて話し合っていますから、「科学と学問」と言うべきですね。SPARC では、オープンスカラーシップについて話し合っています。デジタル・ヒューマニティーズと社会科学では、私たちのキャンパスの教育分野で行われている研究は、完全なアクセシビリティと完全な再利用可能性という戦略の活用の恩恵を受けていると考えています。

オープンアクセス

オープンアクセスの背景についてお話しすると、15年前のブダペスト・オープンアクセス・イニシアチブにおいて、ソロス財団によってあるグループが立ち上げられ、「研究者が研究の成果を完全かつ無料で共有し、人々が最大限活用できるようにするシステムを構想し、構築することができるとしたら、そのシステムはどんなものになるだろうか？　それを実現するには、何が必要だろうか？」という問いに答える任務が課せられました。オープンアクセスという概念は、このブダペスト・オープンアクセス・イニシアチブの開催中に生み出されました。それは、インターネットという新しい技術と、お互いにアイデアや研究結果を自由に共有する科学者や学者の伝統を融合させ、この二つを初めてつなぎ合わせて強力で前例のない公共財を生み出す、科学や個々の科学者だけでなく社会全体に真のプラスの効果をもたらすというものでした。

ブダペスト・オープンアクセス・イニシアチブからの引用をご紹介しましょう。「オープンアクセスが可能にする公共の利益とは、査読付きジャーナル文献を世界中に電子配布し、全ての科学者、学者、教師、学生、その他好奇心旺盛な人々に完全に無料で無制限のアクセスを与えるというものである」。SPARC は、その議論のテーブルにつく特権にあやかり、この声明文の起草を手伝いました。ブダペスト・オープンアクセス・イニシアチブはオープンな研究環境を可能にするための理論的な話だけでなく、そこにたどり着くための枠組みの定義も示しました。

アクセスをオープンにして無料にすると言うとき、人によってオープンという言葉の意味が異なります。資料にオープンにアクセスするには、科学活動や研究を行う必要があります。ブダペスト・オープンアクセス・イニシアチブの定義はとても明解です。「この文献への『オープンアクセス』とは、公共のインターネット上で無料で利用することができ、どんなユーザーでも、金銭的、法的、技術的障壁なく、閲覧、ダウンロード、コピー、配布、印刷、検索、論文の全文へのリンク、インデックス化のためのクロール、ソフトウェアへのデータとしての受け渡し、その他あらゆる合法な目的で使用できるということである」。私たちは、ただ論文へのアクセスを手に入れてウェブ上で PDF として読むだけでなく、オープンなウェブ上のデジタルデータを一括ダウンロードして、そのデータのテキ

ストマイニングやデータマイニングを行い、配布し、インデックス化のためにクロールし、合法な目的のために使用できるようにしたいのです。

オープンアクセスが二つの定義から成ることを覚えておくことが大事です。つまり、無料でオンラインで利用できることと、デジタル環境でその資料を完全に利用する権利です。この定義を完全に実現せずに、オープンサイエンスやオープンリサーチにおける真のオープン性を得ることはできません。SPARC の私たちがこの方向に向けて変化を起こそうと努力する中で本当に重要なことは、オープン化という問題についてどのように話すかということでした。研究者や学者、資金提供者、出版社、科学・研究エコシステムの全ての重要な関係者に、オープンアクセスはいいことだから推進すべきだとか、オープンの方がクローズよりいいとか、私たちはオープンアクセスの賛同者だということをただ言うだけでは、私たちはオープンだということだけしか問題にしていないように聞こえてしまいます。しかしながら、実際のところはそうではありません。私たちは、オープンな方がクローズよりいいからオープンサイエンスシステムやオープンアクセスシステムを求めているのではないということを人々に本当に理解してほしいと思っています。私たちは、誰でも毎日の仕事で重要な進歩を遂げられるようにするための戦略として、このオープン性、無料で利用できること、完全に利用できることを活用したいのです。科学者であろうと、学生であろうと、資金提供者であろうと、私たちは最終的な目標をもっと効率良く達成するための手助けとするために、オープン化に何ができるのか考えなければなりません。

私たちは、ブダペスト・オープンアクセス・イニシアチブが起草されたとき、このことを分かっていました。私たちは、システムの意欲的な目標を設定しました。私たちはそこに到達する方法の枠組み、定義を定め、研究を加速し、教育を豊かにし、富者の学びを貧者と、貧者の学びを富者と共有し、このような文献を可能な限り活用し、共通の知的対話と知識の探求において人類を一つにする礎を築くためのオープン化について話し合いました。こういったことを全て行うためのオープン化であり、クローズよりオープンの方がいいからというものではないのです。

私たちは今、ジカウイルスによる蚊媒介性流行病の蔓延を防ぐためにより素早くデータを共有するといった例のように、オープンな研究環境において資料をオープンにすることになぜキャンパスの研究者が参加すべきなのか、以前より研究者たちと話し合うようになってきています。自分の研究データと論文へのアクセスをオープンにすることは、透明性と再現性の改善にも役立ちます。あなたが話し掛けている人の立場に立つことで、オープン化が彼らにとって大事なあらゆることの達成を可能にする戦略になり得ることを理解してもらえるようになります。

SPARC では、このエコシステムで資料へのアクセスをオープンにすることのあらゆる側面に取り組んでおり、これには教科書も含まれます。教科書へのアクセスをオープンにすれば、どんな学生でも、高等教育にもっと手が届きやすくなります。

「～するためのオープン」

「～するためのオープン」の空白部分を埋めるというアイデアは、グローバルオープンアクセスウィークに参加した方には馴染みのあることです。これは 2017 年のオープンアクセスウィークのテーマでした。私たちはコミュニティに対し、人々に話し掛け、空白を埋めるよう促すことを考えてもらいました。私たちはステークホルダーから、オープンサイエンスやオープンリサーチの企ての最終的な成功にとって何が重要か、聞き出したかったのです。なぜオープン化が重要な戦略になり得るのか認識できるよう取り組んでもらいたいと思ってそうしました。

そして、幾つかの素晴らしい結果が得られました。図 10 は、米国に拠点を置く巨大な民間慈善団体であるゲイツ財団のロビーにある黒板の絵です。オープンアクセスウィークの期間中、ロビーの黒板にこの絵を

描き、資金提供者に「オープンアクセスがあなたのためにできることは何ですか？」という質問に答えてもらいました。ゲイツ財団の人々が実際に空白を埋めてくれました。彼らは世界の保健イニシアチブや公衆衛生に多大な投資を行っています。彼らの「～するためのオープン」はこのようになりました。「ゲイツ財団は、世界の保健イニシアチブへの資金提供を通じて命を救うために、オープンアクセスを支持する」。つまり、彼らは本当にその最終目的を理解していたのです。オープン化を受け入れ、オープン化の力を理解する資金提供者は、オープンリサーチの企てを効果的に実現する非常に重要なピースの一つです。

世界銀行もグローバルオープンアクセスウィークに参加し、このように「～するためのオープン」に記入しました（図 11）。世界銀行のコアミッションの一つは極端な貧困を根絶させることです。世界銀行は過去

（図 10）

（図 11）

5 年間、全ての刊行物とデータへのアクセスをオープンにしています。そのコアミッションを達成する手助けとするため、オープンアクセス方針を採用しました。研究資金の財源をオープンにしたり、人々にオープンであることの重要性を理解させたりするということは、私たちが一生懸命取り組んでいることの一つです。

米国の元副大統領ジョー・バイデンは、オープン化やオープンサイエンスの力を直感的に理解した人物の好例です（図 12）。息子に脳腫瘍が見つかったとき、バイデン副大統領は米国のある治療施設から別の施設に息子を転院させて実験的治療を受けさせようとしたのですが、そのときにオープン化の重要性に気が付きました。一つ目の研究施設が脳の電磁スキャン画像をもう一つの研究施設に転送しなければなりませんでした。そのデータがもう一つの研究施設にはないフォーマットだったため、スキャン画像の読み取りには 24～48 時間かかると、副大統領は言われました。データを変換しなければ読み取れなかったのです。バイデン副大統領は米国癌学会でスピーチを行い、あるフォーマットから別のフォーマットに変換するために必要だったこの 48 時間という途方もない遅れを取り上げました。彼は、相互運用性や、科学データのフォーマットをオープンで簡単かつ即時に読み取れ、完全に再利用できるフォーマットにするための標準規格の使用といったことについて話しているつもりではなかったのですが、直感的にそこを目指していたのです。それから、彼はがん研究者へのスピーチの中で、データが

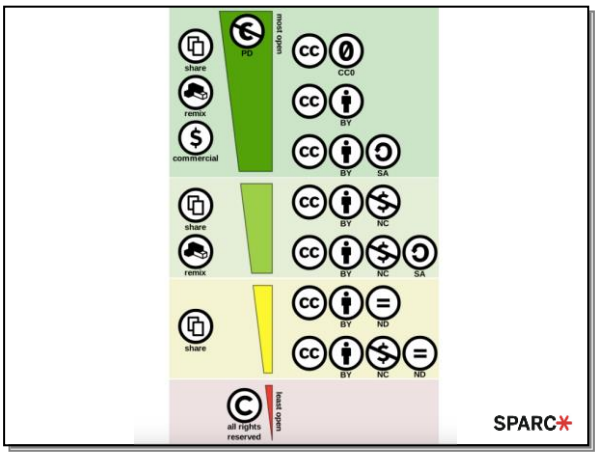
（図 12）

その日のうちに利用できるフォーマットで存在するというだけでなく、そのデータに対する見識を提供する論文にすぐにアクセスできることも大事だと述べました。こうして本質的に、治療の進歩を加速し、結果的に特定の病状や病気に対する治療法を見つけるには、オープンアクセスが必要であるという結論に至ったのです。

他の人の例やその発言に興味がありましたら、openinorder.to というウェブサイトを見てください。オープンアクセスウィークで集めた事例が掲載されていて、オープンサイエンスやオープンリサーチについて話し合うコミュニティの人々にその対話の基礎を成す資料をより多く提供するのに役立ちます。

オープンアクセスへの障壁

私たちは前進しています。オープンアクセスは手の届くところにあります。しかし、私たちが 15 年間もこのことを話してきたのに、いまだに実現していないのはなぜなのか考えることが重要だと私は思っています。商業出版社のせいだという人もいるでしょう。商業出版社の抵抗はある程度ありますが、一番の問題ではありません。一番の問題はオープンサイエンスやオープンリサーチ環境への移行が最適だと納得していない、あるいは理解していない個々の学者や研究者です。これは驚くべきことではないでしょう。論文をオープンに共有する選択肢があると分かっていても、大抵の場合、あるジャーナルが本当にオープンアクセスなのかどうか、見極められません。クリエイティブ・コモンズのライセンス体系は非常に分かりにくいものです（図 13）。私たちは研究者たちに、本当にオープンなのかどうか見極めるために、簡潔にした法律用語の勉強をお願いしています。



（図 13）

論文を掲載するために、論文掲載料を徴収して著者に手数料を払わせるハゲタカ出版社が急増しています。私たちは研究者から、偽物だと分かっているジャーナルに論文を掲載するよう求めるメールが週に 5 通も 10 通も届くと聞かされています。そのせいで研究者は、オープンアクセスはいいかげんで、合法でなく、オープンアクセスジャーナルに掲載するのは良くないことだと思ってしまいます。研究者は自分のデータをどうすればいいのか混乱しています。私たちは、どうすればいいのか説明して、そのデータをオープンに利用できるようにしたり、シェアできるようにしたりする方法を理解するのに役立つメカニズムを提供しようとしていますが、それでも混乱は解消されません。理解しようとしなければならない細かなことがとにかくたくさんあるのです。

最後に、私たちが本当に進歩を遂げるためにしなければならない三つのことにフォーカスを当てたいと思います。研究者が一番気掛かりだと私たちに言ってくることは、研究成果をオープンに利用できるようにしても、終身在職権、昇進、資金、評価の面で、高いインパクトファクターがあるジャーナルに掲載したときのように見返りが得られないということです。私たちはこの問題に取り組み、真正面から対処しなければなりません。オープンな資料やプラクティスを作って活用することが報われない制度上のインセンティブがあるということです。このような制度上のインセンティブは大学のキャンパスでも、官民の資金提供組織でも見られます。

オープンアクセスへの障壁への対処

最後に、研究者が直面している問題、そしてオープン環境への移行を目指す私たちが直面している問題への対処方法として、三つの提案をしたいと思います。第一に、私たちは科学・学術のコミュニケーションプラクティスを支えるインセンティブを再調整し、大学のコアミッションに貢献する人々に報いるようにしなければなりません。大学や資金提供組織のミッションステートメントを読むと、この千葉大学の例のように、知識の普及がその方向性としてよく掲げられています（図 14）。この事実にもかかわらず、私たちのインセンティブでは、できるだけ幅広い読者に届く訳ではない媒体への発表が報われるようになっています。従って、インセンティブと報酬の行動を再調整し、大学や資金提供者がそのコアガイドラインの方向に向かうようにしなければなりません。

米国などには終身在職権や昇進のガイドラインがとにかくたくさんあります（図 15）。最初の列には、平均インパクトファクターが何点以上のジャーナル 10～15 誌に掲載されると、教授は報奨を受けると書いてあります。到達する読者の幅や、デジタル環境でのその成果の実用性には触れていません。これは変えなければいけません。私たちは、オープン化をもっとサポートするために、インセンティブを作り直し、再調整することを考える必要があります。

インセンティブは、学問領域によって異なります。オープンへの道筋は、遺伝学やゲノミクスと他の領域では異なるでしょう。皆さんの学問領域ではこういった報酬や行動をどのようなものにできるか、考えてみてください。

私たちは、あらゆる種類の個々の利用と潜在的影響を追跡する論文レベルのメトリクスの使用を検討し、これを評価プロセスに取り込むことができます。インパクトファクターを他のメトリクスに置き換えようと言っているのではなく、メトリクスで補完するということです。私たちは、資料をオープンにシェアする行動に報酬を与える追加の尺度を導入すべきです。

米国では、連邦政府が科学の重要性を認識し、科学への資金提供の継続に取り組む中、民間財団や民間資金提供者などの研究資金提供者はますます重要な役割を果たすようになってきています。米国の民間財団は研究が継続できるようキャンパスに多額の資金を提供し、ゲイツ財団のように、大学がその第一のミッションを果たせるようにするためのコア戦略として、オープン化のアイデアを採用しはじめています。

Open Research Funders Group は、2 年前から SPARC が支援している資金提供者のグループで、幾つかの米国最大級の民間資金提供者が参加しています（図 16）。同グループは、研究者を支援していると宣言するための方法を模索しており、投資益を高めることを目指しています。研究に関連する結果、データ、論文へのアクセスをオープンにすることで、そのミッションの達成を支援できると考えています。また、オープン化が一つの選択肢以上のものになるようにインセンティブ

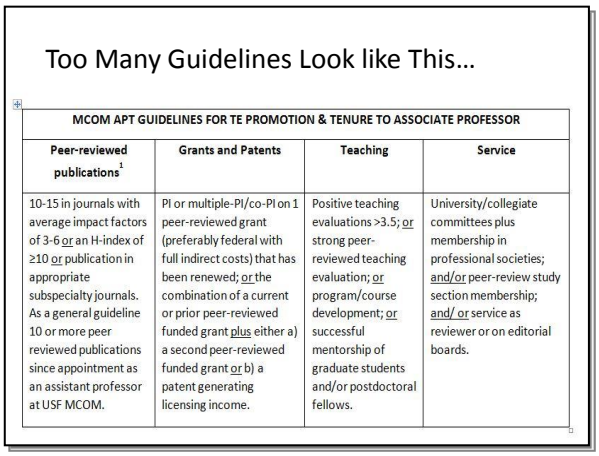
Home > About Chiba U > Message from the President/Vision

Chiba University began with five faculties in 1949 as a national university under the new postwar education system. Since then, it has expanded to become a comprehensive university comprised of ten faculties and 11 graduate schools.

Since its foundation, Chiba University has aimed to contribute to society through its pioneering education and research. To attain this goal, faculty and staff are encouraged to think outside the box and to collaborate across specialty and profession. This open system of faculty-level exchange enables Chiba University to promote its distinctive integrated education and research.

In 2004, changes to the education law allowed the national university to become an independent administrative entity. Under our university's motto, "Always Aim Higher", we established a new goal to cultivate human resources capability of assuming roles in global society. In 2016, we created a new faculty, the College of Liberal Arts and Sciences, in an effort to stimulate the further development of our university mission.

（図 14）


Too Many Guidelines Look like This…

MCOM APT GUIDELINES FOR TE PROMOTION & TENURE TO ASSOCIATE PROFESSOR

Peer-reviewed publications[1]	Grants and Patents	Teaching	Service
10-15 in journals with average impact factors of 3-6 or an H-index of ≥10 or publication in appropriate subspecialty journals. As a general guideline 10 or more peer reviewed publications since appointment as an assistant professor at USF MCOM.	PI or multiple-PI/co-PI on 1 peer-reviewed grant (preferably federal with full indirect costs) that has been renewed; or the combination of a current or prior peer-reviewed funded grant plus either a) a second peer-reviewed funded grant or b) a patent generating licensing income.	Positive teaching evaluations >3.5; or strong peer-reviewed teaching evaluation; or program/course development; or successful mentorship of graduate students and/or postdoctoral fellows.	University/collegiate committees plus membership in professional societies; and/or peer-review study section membership; and/ or service as reviewer or on editorial boards.

（図 15）

を再調整することを、高等教育の関係者と共に検討することを求めています。これは、オープン化に向けた変化を開始するための非常に有望な戦略です。

次に、私たちはオープンアクセス出版を支援し、全体的なオープンアクセスでの発表数を増やすビジネスモデルを考え直す必要があります。学問や科学が世界的な事業であることを認識することが本当に重要です。今日の最初のお二人のお話は、研究の全体的な性質とオープンな共有、特にジャーナルの共有を支える財務モデルを構築する必要性をまさに描き出してくれました。また、このようなモデルがローカルな知識の創造を支援し、小規模な研究機関やあらゆる種類の寄付者の発展途上国での貢献を促し、全世界への知識の普及に報いるようにしなければならないことを示してくれました。

近年、オープンアクセスコミュニティでは、購読ベースのジャーナルから論文加工料ベースのジャーナルへの大規模な移行について考える動きがあります。これはオープンアクセスの論文がもっと増えることにつながる興味深いモデルですが、私たちは、そうすることが購読料を払えずに科学へのアクセスを拒絶されている国に与える真の影響について考えなければなりません。こういった国々には、オープンアクセスジャーナルに論文を掲載するためのお金もありません。従って、科学への参加の問題が出てきます。私たちは、オープンリサーチのコミュニケーションチャンネルを支えるビジネスモデルの公平さや公正さについて考えなければなりません。

（図 16）

三つ目、最後のトピックは、同じくジャーナルでの発表をサポートするチャンネルにおける公平さとオープン性の精神にのっとったものですが、学術研究環境で使用するインフラを広い範囲でオープンにし、「コミュニティ所有」や「コミュニティ管理インフラ」と呼ばれるものにしなければならないということです。私は、この点に関して機関リポジトリを特に念頭に置いています。エルゼビアは最近、機関リポジトリプラットフォームの bepress を買収しました。bepress は営利企業で別の営利企業に買収されたということですから、それ自体は問題ありません。この買収は図書館・研究コミュニティにとって驚くべきことではありませんでしたが、図書館がキャンパスで論文を無料かつオープンにシェアするために頼りにしているプラットフォームがもう私たちのものではなくなってしまったということに、多くのライブラリアンの頭の中でアラームが鳴りました。このプラットフォームは営利企業に所有されています。それは問題になるかもしれないし、ならないかもしれませんが、このチャンネルを自分たちが所有しているということは、ライブラリーコミュニティでは当然のこととして捉えられていました。私たちは、このプラットフォームがいつまでもオープンな世界の一部であり続けると思っていたのです。私たちはインフラの営利ベンダーと交渉するときには、条件を巡って賢く振舞い、たとえ所有権が営利企業の手に移っても、そのインフラの所有権とオープン性には手が出されないよう保証しなければなりません。私たちはどこかに向かっているのではありません。私たちは景色の一部になろうとしているのであり、営利企業と一緒に働くためには、もっと賢くなる必要があります。

最後に、そして最も重要なことですが、私たちは本当に、研究事業の全体的な性質を反映した共同戦略の策定を継続する必要があります。関係者全員のために研究を共有するシステムを最適化するため、ゼロからシステムを考え直すことができるとしたら、どんなも

のになるでしょうか？　私たちはこの質問を自問し続け、一緒に答えを探し続けなければなりません。それから、オープン化を私たちが共有する目標の達成を実現する効果的な戦略として活用するシステムの構築に取り組まなければなりません。

●武田　プレプリントサーバには長い歴史がありますが、10 年前はまだとても限定的な領域でした。今、プレプリントはものすごい勢いで成長しています。アクセス環境をコントロールし続けなければならないとおっしゃっていましたが、コミュニティが管理しているプレプリントサーバとそうでないものがあることを考えると、オープンアクセス環境においてプレプリントサーバはどのように位置付けられるべきでしょうか。

●Joseph　誰がプレプリントサーバをコントロールしているか、所有しているかという点では雑多な状況になっています。私は、ID はインフラだという武田先生のアドバイスに従わなければならないと思います。プレプリント向けの DOI が必要です。ジャーナルに掲載された論文や、機関リポジトリにあって DOI を持っている論文と同じように資料の位置を定められるよう、著者に ORCID 識別子をつける必要があります。インフラの重要ピースとしての識別子という概念は、プレプリントサーバを成功させる要因の核心を突くものです。

●武田　つまり、プレプリントサーバはオープンアクセス環境を整える上で重要な要素の一つだということですね。

●Joseph　もちろんです。そして、プレプリントサーバの使用が増加しているということはとても励みになることです。物理学では以前からプレプリントサーバを使っていましたが、ここ 2 年ぐらいで、プレプリントサーバはついに、これは使えるかもしれないとみんなが思うような姿になりました。行動の変化が起こっていて、研究者たちは、プレプリントは自分たちが研究をもっと早く進める上で有益だと考えるようになってきました。研究者たちは、プレプリントサーバに何かを載せてもすぐには見返りが得られないかもしれないけれど、そうすることで被害を受けることもないということが分かってきました。これは、コミュニティでの行動の変化を促す上で本当に重要な心理的ステップです。

●David　「～するためのオープン」というアイデアの話に戻りますが、あらゆる活動のフォーカスは実質的に外部性の創出に還元することができます。外部性というのは、化石燃料を燃やすといった悪いこともありますが、何かを公開して他の人が利用できるようにしたり、きれいな水を利用できるようにしたりするといった良いこともあります。私たちは外部性の影響を大いに受けますが、これは従来の資源配分方法の計算の範囲内には収まりません。従って、経済学には、良い外部性を提供しようとすることを扱った文献や、何世代にもわたる悪い外部性を止める方法を扱った文献が大量に存在することを知っておくことが重要です。これは取り組みやすい分野ではありませんが、いろいろな人に働き掛けて、経済学で外部性について学んだ指導的立場の人やもっと若い人たちを見つけ、多くの人々にとって有益な行動を引き出し、何らかの見返りを得られるようにするために使われてきたメカニズムを明らかにすることができると思います。これにより、小さなグループにとってしか利益にならない外部性に対しても貢献するよう人々に促すことができます。問題は、とても大きなグループにとっての外部性を生み出すような方法で小さなグループに協力させることにより、小さなグループの活動を生かせるかということです。これは興味深いアプローチで、実際にこのようなアプローチが取られてきました。例えば、オゾンホールの解消に使われるメカニズムを考えることで、人々は気候変動の問題に取り組んできました。

●Joseph　これは大事なポイントですね。ここから指摘しておきたいポイントが二つあります。私たちはこういったことを実証する方法について考えていて、また、SPARC は若手の駆け出しの研究者と一緒に取り組むという考えを特に重視しているということです。私たちは世界中の若手研究者を集めて毎年 OpenCon を開催しています。毎年だいたい 200 人参加しています。昨年は世界中から 1 万 3,000 人の若手研究者の応募がありました。そのうち 200 人しか受け入れられないのですが、62 カ国から受け入れ、全員が会議に出席しました。この会議の活動では、個人またはグループで取り組みたいこと、確認したい共通性を特定し、活用して、前進を可能にするための戦略としてオープンにすることが含まれていました。今では、この会議の参加者の中から、それぞれの専門領域や関心分野で進歩を遂げる方法としてオープン化を活用するコミュニティを形成する若手研究者のグループが成長しています。つまり、これは始まりだと思っています。私たちはこれを経済理論の基礎を活用するものだとは考えていませんでした。これを実践コミュニティとみなし、こういう方法でコミュニティ活動を構築していこうとしていたのです。こういうふうに考えることは非常に重要で実りの多いものですね。

●村山　私の質問は、今日のお話の核心に触れるものではないかもしれませんが、さまざまな国でのオープンアクセスの実践の上で、国によって出版業界に異なる文化があったり、異なる行動があったりするか、それは司法制度の違いによるものかもしれませんが、これまでの経験の中でそういったことに気付かれたかどうか、関心があります。私はアジアにだけ焦点を当てているのですが、こういった経験をお持ちでしたら教えてください。

●Joseph　商業出版社に事業の仕方を教えるようなつもりはありません。商業出版社はとても賢く、さまざまな国でさまざまな経営方法を探っていると思います。私たちの観点からすると、私たちはそれぞれの地域でうまくいくさまざまなモデルを模索しています。例えば、論文加工料に代わるものを検討する中で、機関投資家による資金提供を通じてローカルコンテンツを支援するラテンアメリカの文化について考察しています。ラテンアメリカには、大規模な商業ジャーナルに発表したり、インパクトファクターを重視したりする文化がありません。多くの国に、地元で生まれたコンテンツに価値を置く文化があるからです。ジャーナル名ではなく大学の出版許可が重要です。私たちはこういった文化の違いに目を向けています。

私たちはまた、高いインパクトを持つジャーナルへの掲載が求められるため、ローカルな知識の創造が同じように評価されていない場所についても注意深く追っています。私たちは最近、同僚たちと一緒にアフリカに SPARC の事務所を立ち上げました。多くの国で、研究者は一次言語として英語を使用する商業ジャーナルやインパクトの大きいジャーナルでの掲載が求められていて、駆け出しの研究者にとって不利になっています。次に、こういうジャーナルの関心事となる主題は、多くのローカル研究者にとって重要な主題とは大きく異なっています。土着の、ローカルな研究に関心がないのです。アフリカの同僚たちは、学術コミュニケーション・出版の脱植民地化を求めています。

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

デジタル時代の研究プロセスと大学、大学図書館における支援のあり方

倉田 敬子

（慶應義塾大学文学部）

講演要旨

多様な情報源の探索、研究データの収集、分析、成果の発表まで、研究者たちの研究プロセスはデジタル化されつつある。大学図書館は従来、研究に必要とされる資料を提供することでその研究活動を支えてきた。学術コミュニケーションにおけるデジタル化とオープンアクセスの進展は、大学図書館のこの役割を縮小させているといえる。大学は大学図書館の予算および研究者たちの経済的基盤を提供することにより、研究活動を支えてきたが、グローバルな競争とオープンサイエンスを志向する国の政策や社会状況の中でより直接的な関与がもとめられてきている。大学図書館、さらには大学というコミュニティは研究者の研究活動をいかに支援できるのか、研究プロセスのデジタル化という観点から考えてみたい。

倉田 敬子

1987年慶應義塾大学文学研究科博士課程修了。1988年慶應義塾大学文学部図書館・情報学科助手、1993年同助教授、2001年から文学部図書館・情報学専攻教授。専門は学術コミュニケーション、特に情報メディアのデジタル化、オープンアクセス、研究データに関心を持っている。著者に『学術情報流通とオープンアクセス』（勁草書房）など。2008年度日本図書館情報学会賞授賞、2010年科学技術への顕著な貢献2010（ナイスステップな研究者）選定。SPARC Japan 運営委員会委員。

オープンサイエンスとは何なのか

今回のセミナー全体としては、オープンサイエンスとは何かが大きなテーマです。特に午前中は、科学そのものとは何なのか、データドリブンとはどういうことなのかということが話の中心でした。でも、私はやはりオープンサイエンスとは何かはよく分かりません。データが重要だということは、研究者なら誰だって分かっている話ですが、データドリブンだと言われてしまうと、今までの科学とは何が違うのかということになります。いわゆる第 4 の革命というような、データがまずあって研究が始まるというところまで本当にいくのかは、社会科学、人文科学の立場から言うと、まだよく分からないし、制度化という点ではまだまだ未知数のところがあると考えています。

オープンサイエンスのもう一つの側面は、デジタルであるということです。非常に広い文脈で社会全体がデジタル化している中で、研究も完全にデジタルの世界へ行くだろうといわれています。

もう一つ、個人での研究がないわけではもちろんありませんし、個々の研究自体はもちろん研究者個人がするのですが、プロジェクトとしては、共同・共有・協働が科学の世界においても非常に重要視されています。これまで個人が中心だったといわれていた人文科学も含めて、共同・共有・協働が研究の一つの特徴と

されるようになっていると考えています。

このデジタルと共同（協働）という環境の中で、今、研究者たちはどのような研究をしているのか、こちらの方に今日は焦点を当てて事例を紹介いたします。このようなことを「デジタルスカラーシップ」と言う方もいると思います。ただ、今回はオープンサイエンスがテーマなので、これもオープンサイエンスの一つの側面と考えます。

デジタル時代の研究者

いろいろな焦点の当て方があると思うのですが、私は「研究ライフサイクル」という表現に関心を持っています。最近よく目に付くのでググってみたところ、図 1 のような図がたくさん出てきました。Pinterest でもたくさんピンされており、面白かったので幾つかここに示しています。後ろに隠れている、文章だけで表しているようなものもありますが、やはり写真映え、インスタ映えするような図にしている方が非常に多くて、図書館のサイトにもこのような図を使って図書館でのサービスを説明しており、すごいなと思った次第です。

研究ライフサイクルの各要素を見ると、共通している要素がかなりあります。図 2 は、Kramer と Bosman が 2015 年のカンファレンスで発表したものです。学術コミュニケーションの 101 の新機軸、イノベーションと言っています。内側の円に 2005、真ん中の円に 2010 と書いてあります。一番外側が 2015 年という意味なのだと思われます。ポスターなので細かい発表は分からないのですが、これはデジタルに焦点を当てたもので、このような形で学術コミュニケーションを支えるさまざまなサービスが出てきているということです。

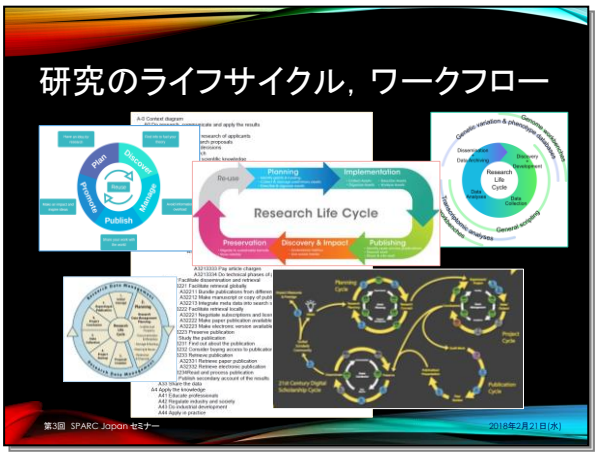
それぞれのワークフロー、研究のライフサイクルに応じてこれら新機軸のサービスを利用していることを示しています。最初の段階が Discovery（発見）です。Web of Science のようなデータベースを使っていた人たちが、今や Google や Google Scholar、あるいは arXiv のようなプレプリント・サーバで論文・文献を探すようになっています。

Analysis（分析）に関しては、SPSS など個々のソフトウェアを使っていた人たちが、今やインターネット上で盛んに共有できるようになっている R や GitHub を使いだしています。

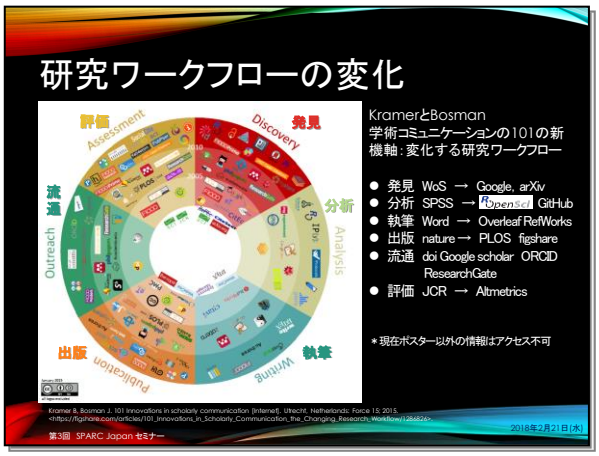
Writing（執筆）は Word になって長年たちますが、Overleaf や RefWorks をみんな使いだしているのではないかということです。

Publication（出版）は申し上げるまでもなく、伝統的な「Nature」のような雑誌から、オープンアクセス・ジャーナル、さらには figshare で学会発表のものを共有するということも出てきています。

Outreach（流通）は、もちろん DOI、ORCID といった ID を付けることから、Google Scholar でも業績の一覧を作れますし、ResearchGate で論文を共有することもあります。この Outreach（流通）と Discovery



（図 1）



（図 2）

（発見）のところはある意味では表裏だと思うのですが、いろいろな形で情報が手に入るようになっています。

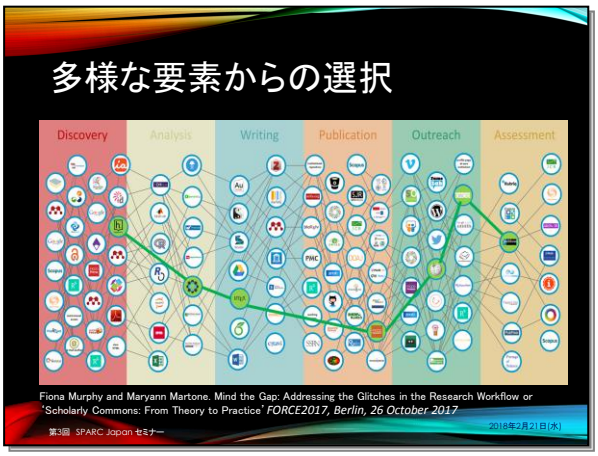
最後は Assessment（評価）です。以前は JCR やインパクトファクターぐらいしかなかったものが、altmetrics などさまざまなものが出はじめています。このように、現在、研究のワークフローにいろいろな変化が出てきているということが言われています。

図 3 はさらに 2017 年に FORCE2017 という学会で発表されたもので、今、研究者は、それぞれの段階でさまざまな新機軸、サービスの中から好きなものを選べる段階に来ているということです。これは逆に、選ばなければいけなくなっているということで、何をどういう形で選んでどうやっていくのかを決めるのは、これはこれで結構大変です。

一人の研究者として

ここまでが全体としての概略で、では私はどうしているのかを簡単にお話しします。本当であれば、研究者は今このように変わっていますと言いたいのですが、これはまださまざまで、それについての調査などの成果もそれほど出ていないので、社会科学者としての私の一例をご紹介します。

大学の研究者の場合、複数の目的・役割を同時進行で進めていかなければいけません。もちろん教育にも非常に時間を割かれます。教育は自分の研究が土台になりますし、大学の研究者には研究者の育成という役割もあるので、これをないがしろにはできません。でも、自分の研究と教育はやはり少し違います。



（図 3）

次に、大学や学会において何らかの役割を果たさなければならず、デジタル化の影響はここにも来ています。研究を応用できる側面ももちろんなくはないのですが、それとは全く違うことが求められます。

さらに、研究プロジェクトが一つという方もそうはいらっしゃらないと思います。やはり幾つも同時進行していて、目的もステップも異なります。例えば、このような講演を頼まれるということも入ってくるわけです。そうなったときに、どういう形で、自分で自分の行動をコントロールしていくのかということが非常に難しくなってきます。

発見の側面もデジタル化されて、情報探索の側面はほぼ完全にデジタル化されていると思いますが、どれをどう使うかは組み合わせていると思います。私もいまだに図書館の OPAC はよく使いますし、その一方でやはり Google Scholar は手離せませんし、ResearchGate からもメールはよく来るし、何かあれば figshare をちょっと覗いてみる、それを当たり前のようにもうやっているわけです。

しかし、その中で、雑誌そのものを見に行くということもやはり今も行っています。重要な雑誌のモニタリングをやめるということはしません。そういう最新の情報に触れるというカレントアウェアネスと、研究プロジェクトのそれぞれの段階に合ったところで網羅的に探す、この二つが似ているようでいて、やはり全く目的が違うので、文献、情報の探し方も違ってくるということになります。

さらに、これは当たり前なのですが、今、入手する段階において、デジタルで入手できるという点では非常に便利になりましたが、残念ながらまだ図書や日本語の論文に関しては紙でしか手に入らないものがあります。読むときには、私はまだ紙が優位で、デジタルも増加はしているのですが、紙で読む場合がまだまだあります。何かを記録しようとするときのアイデアのメモや文献の注釈はデジタルで随分できるようにはな

ってきていますが、やはり紙も使っています。読んで、情報を整理して、メモを作ってという段階を、私はもっとデジタルで統合化したいと常々思っているのですが、これは結構難しいのです。さらに、共同研究になると、ここが非常に複雑になるので、この辺でもう少し効率的なサービスができないのだろうかということは常々思っています。

次の段階で、調査を計画しようとした場合、もちろん先行研究を集めてきて参考にします。先行研究でのデータも最近は入手可能ですが、これまでも例えば社会調査の質問票などは情報として十分参考にしていました。でも、これからは個票データが入手可能になります。現にもう幾つかの社会調査ではデータまで公開しています。では、それで調査計画のやり方が変わるのかというのは、私にはまだよく分かりません。理論があって目的があって、自分なりの考えで調査を計画しています。そのときに人のものは参考にはなるけれど、そこでデータが手に入ったということで、本当に研究のやり方まで変えるのかというのはよく分かりません。

さらに大変なのは助成金を取ることで、今は、学内でも調査は倫理委員会にかけなくてはなりません。また、日本ではまだ大きくなってはいませんが、欧米でいわれているデータ・マネジメント・プラン（DMP）を今後は出していくことになると思います。煩雑な書類や手続きが必要で、これはできれば、研究者ではなく、もっと違う人たちに助けてほしいとは思います。ただ、それもやはり研究の一部なわけで、どこまでをどのように研究者としてうまく効率化しながら、やらなければいけないことをきちんとやっていけるかというのはなかなか大変であるというところに来ていると思います。

私自身は調査の分析で、分析ツールまで共有することはまだそれほど多くはないのですが、やはり統計や、例えば AI とまでいかなくても、そのような最新の研究をされている方の中では、分析ツール自体を共有する、分析の方法自体を共有することが始まっています。このような傾向をどのように自分の研究の中でうまく管理していくのかというのは、非常に大変になってきていると思います。

そんなに大きな調査ではなくても、少し調査をすればデータは山のように出てきて、解析はそのたびに、段階ごとに無数のファイルを作ってしまうので、これをきちんと管理するのはとても大変なことです。基本的に今、私は Dropbox や Wiki、Dropbox Paper などを使っていて、最後のデータは仕方がないので Zenodo や figshare で公開するという方法を取っています。

執筆・発表の段階になると、これはある意味ではこれまでの伝統的なやり方に従うのですが、基本的には共同研究などでは完全に共同執筆になります。部分ごとに分担して書いて、それを上げて、みんながもう一度全部読み直して、人が勝手に手を入れていってということを何回もやることで論文が出来上がっていきます。

最後は、やはり英文校閲にかけなければいけないので、ここは業者にお願いすることになります。今は引用の形も、場合によっては投稿まで業者が引き受けてくれます。これは研究者にとってはとても便利で、頼みたくなります。というのは、投稿はオンラインシステムなのですが、雑誌によってまちまちで、そのたびにやり方が少し違うため、すごく面倒なのです。それを例えば 1 万円でやってくれると聞くと、ついふらふらと頼みたくなるという気持ちは大変よく分かります。

Outreach（流通）のところは、一番変わっているのではないかと思います。今まで研究者自身が outreach を意識することはあまりなかったと思います。でも、例えば今、私は Google Scholar にプロフィールを登録していて、そうすると、自分で論文を刊行すると、1 週間ぐらいたつと、Google から「これはあなたの発表ですか」とメールが来ます。それにただ「Yes」とするだけでいいのです。「引用されましたよ」ということもすぐに知らせてくれます。

もちろん ResearchGate にも登録していて、そうすると ResearchGate で「この人があなたの論文を読みまし

たよ」ということを教えてくれます。今まで、そのようにリアルタイムで反応が分かるということはあまりなかったので、そういう意味では、研究者も outreach を直に感じられるようになってきているのではないかと思います。また、ORCID 番号を取っているのですが、これは Scopus と連動してくれているので、ここでも自分の業績がある程度自動的に積み重なっていきます。

それに比べて、これは慶應義塾大学の例ですが、大学の研究成果データベースは非常に入力が面倒で、多くの研究者ができていないと思います。何とか最低限やっているのですが、ここ 1～2 年はそれすら怠りがちです。Google Scholar や ResearchGate のサービスに慣れてしまうと、今の大学のシステムはとても遅れていると感じざるを得ません。ただ、慶應のシステムは新しくなるそうなので期待しています。

大学、大学図書館による支援

では、このような研究者を大学や大学図書館はどう支援できるのかということが問題なのだと思います。前半で、村山先生が学会には学会のやるべきことがあるとおっしゃったのはまさにそのとおりで、学会こそ社会と研究をつないでくれるのです。では大学はとなると、ここはとても難しいところだと思います。学会はアカデミックなコミュニティとして確実にあります。でも、大学はインスティテューションでしかないのです。同じインスティテューションに所属しているのですが、それぞれ違う学会に所属している研究者たちを取りまとめているわけです。その大学の中の図書館に何ができるのかということになると、これはなかなか難しいと思います。

日本ではまだなかなかできていませんが、欧米の大学図書館での研究支援の例は随分たくさん報告されています。ただ、全体として、最後の結論だけ言ってしまうと、その研究支援はインストラクションやコンサルタントなのです。それが悪いとはもちろん言いません。とても重要な点です。例えば、助成の条件がどうなっているのか、データ・マネジメント・プランをどう作ればいいのか、そういうことに関して図書館員がアドバイスしてくれるというのは大変重要だと思います。ただ、これはインストラクションです。要するに、全体の動向をきちんと把握して、それを研究者に仲介する役割だと思います。これで止まっていいのかということが私の疑問なのです。

そうではなくて、学術コミュニケーションシステムを改革するところに大学図書館の役割があるのではないかという話です。例えば、機関リポジトリはそのために作られて、今、それは一つの役割を果たしているとは思いますが、オープンアクセスの中でものすごく重要なステークホルダーになり得ているかというと、残念ながらそこには至っていないと言わざるを得ないと思います。

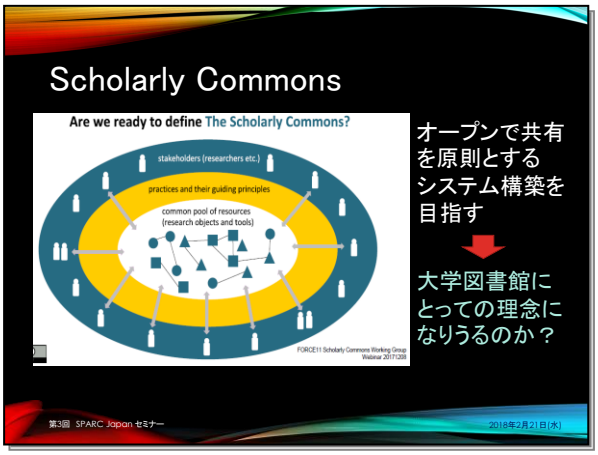
例えばデータ・マネジメント・プランの作成というところでは一定の役割を果たしていますが、データそのものの管理を本当に図書館ができるのかというところは難しいですし、図書館でなくても、大学ができるのかということすら難しいのです。figshare や Zenodo と争って勝てるのかという話です。既存のデータアーカイブとどういう関係を作っていくのかということです。競わなくてもいいのですが、役割分担は必要だと思います。そういう意味で、学術コミュニケーションの生態系（エコシステム）の中で、大学図書館は一体何ができるのかを考える必要があるのではないでしょうか。

前半の David のお話からすると、実現は困難ということになりますがが、全く違うシステムを構築するという運動があります。今、面白いと思っているのは、Scholarly Commons というムーブメントです（図 4）。これは、研究のさまざまな素材やツールなどを全てオープンにして共有するという原則に賛同、協力してくれる研究者、組織、出版社などだけのコミュニティを作れないかというものです。完全にオープンでなくても、ミクスチャーのレベルをできるだけオープンの方に近づけた形で、しかも研究成果だけでなく研究プロ

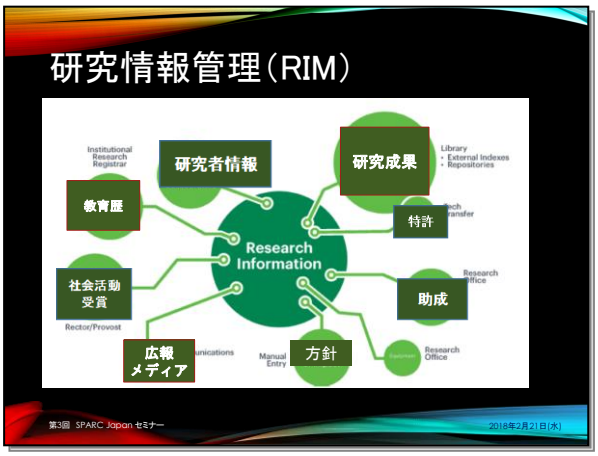
セスのところまでも含んだ形でのコモンズを作れないかという動きです。これはあくまでも一つの理念ですが、そういうところに一体、大学図書館は本当にどこまで関わり得るのかという話だと思います。

そこまでは無理なら、大学でせめてこれくらいできないかなというのが、研究情報管理です（図 5）。大学というインスティテューションの中で、さまざまな研究に関わる情報が生まれ、管理する必要が生じているわけですが、これらが全く統合化されていません。せめて、研究成果、研究者の情報、教育歴、広報、研究助成といった情報の統合的なシステムを作るということぐらいは、大学がやるべきではないかと思いますし、十分やれることではないかと考えています。

図 6 は、Horizon 2020 の一つとして行われている Common European Research Information Format というものです。これは図書館が今まで取り組んできた文献に対するメタデータを、研究情報の標準化にも生かしていこうという考え方だと私は取っています。図書館が扱ってきたメタデータを、もう少し違う形でも生かしていけるというのは、図書館にとって親和性が高い活動として、今後このような形での貢献も考えられるのではないでしょうか。インストラクションやコンサルタントだけではなく、このような形で学術コミュニケーションのインフラのところに関わるような活動を、図書館が行えるなら大きな貢献と言えるのではないかと考えています。

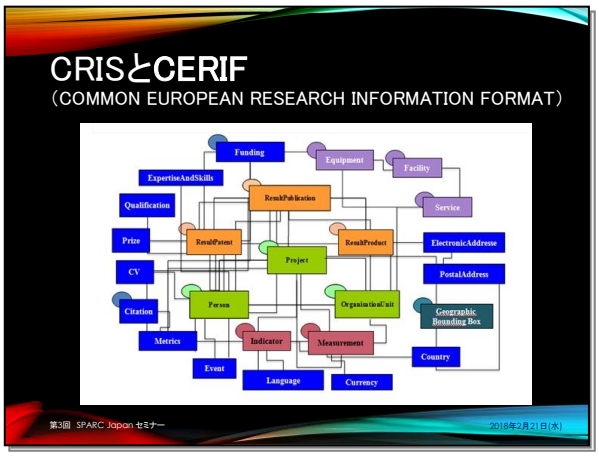


（図 4）



（図 5）

●Joseph　大変素晴らしかったです。Force11 の Scholarly Commons モデルを取り上げてくださってとてもうれしいです。図書館コミュニティでは、中央部分に位置する共有するツールと、運営に関する共通原則について、取り組むべきことがたくさんあります（図 4）。私たちは、こういった原則を定義し、ここでの成果、ワークフロー、ワークスペースをよりしっかりと管理できるようにならなければなりません。



（図 6）

第 3 回 SPARC Japan セミナー2017

「オープンサイエンスを超えて」

デジタル化時代の研究者のために 図書館が構築すべき学術情報環境

市古 みどり

（慶應義塾大学三田メディアセンター）

講演要旨

日本の大学図書館の研究支援は主に資料の収集と提供であった。デジタル化が進んだことにより研究者の情報環境は大きく変化し、図書館の業務および研究支援に対する態度も変化がみられる。しかし、オープンサイエンスを図書館が支援するまでには至っていない。図書館、研究者、出版社の現状について経験から概観し、日本の大学図書館が当面進むべき方向性について考えてみたい。

市古 みどり

三田メディアセンター事務長。信濃町（医学）、理工学および日吉メディアセンター事務長を経て現職。JUSTICE運営委員会委員長。SPARC Japan運営委員会委員。

今日、David 先生のお話を伺って、現場で起きていたことはこういうことだったのかと思いました。また、日本では今、図書館の中で「オープン」というのは何となくキーワードにはなってきているのですが、やはり何をしていいか分からない、でも波は来ているという状況です。Heather さんのお話を伺って、日本の図書館界でも、Heather さんのようにアドボカシーをする人は非常に大事だと強く思いました。そこでまずは私が観察した環境の話からスタートさせていただきたいと思います。

私は慶應義塾大学三田メディアセンターに勤めています。慶應義塾大学ではこの「メディアセンター」という名前が図書館の組織名になっています。「メディアセンター」は、北米の方が聞くと、オーディオビジュアルのことだろうかと思うような、変な慶應発英語名になっているのですが、1993 年に慶應義塾大学の図書館は組織名を「研究・教育情報センター」から「メディアセンター」に変えたのです。このとき私はあまり理解できなかったのですが、図書や雑誌の時代から違う時代へ移るということを見越して、和製英語の組織名に変えたのだと今では理解しています。

図書館がやってきたこと

私たち図書館員は普通に資料を集めて、保存し、提供する、ずっとそういうことを繰り返してきた者たちです（図 1）。これは知識の蓄積や学問の発達に欠かせないものとして私は理解していますし、やってきた仕事は決して間違いではなかったと思っています。メ

ディアが変わったとしても、同じ仕事を図書館員はし続けるのだろうと理解しています。

図書館員が利用者（研究者）に対して大事にしてきたマインド、柱として今でも残っているものは、Ranganathan（ランガナータン）が言った五つの法則です。これは図書館（員）の指針や行動規範になっている気がします。ランガナータンが言った 1931 年の時点では、本という資料に注目していました。けれども、時代が変わって IT、デジタルの時代になり、今、図書館員の行動に望まれるものは、利用者の時間をよく考えることに移ってきたという研究を、Online Computer Library Center, Inc.（OCLC）が発表しています。

K 大学の研究者の情報環境

今から K 大学研究者の情報環境についてお話しします。主に科学、技術、医学（STM）系の話になってしまうと思います（図 2）。K 大学は 17 億円の予算を持っていて、その半分以上を電子資料に費やしているため、大手出版社の電子ジャーナルはもちろんほぼ全て利用可能で、世界のどこの図書館にも引けを取らないぐらいのコレクションを利用できる環境にあります。サブスクリプションしているジャーナルもありますし、オープン・アクセス・ジャーナルもありますし、いろいろなジャーナルがあります。プレプリントサーバを利用されている方もいますし、紀要も利用されています。紀要はいまだに人文科学系を中心として、日本では大きな論文発表のメディアとなっています。

では、検索はどうでしょうか。もちろん大きなデータベース会社あるいは出版社のデータベースを利用できますが、Google や PubMed など有名で無料のデータベースが実は一番よく使われているのではないかという状況になっています。

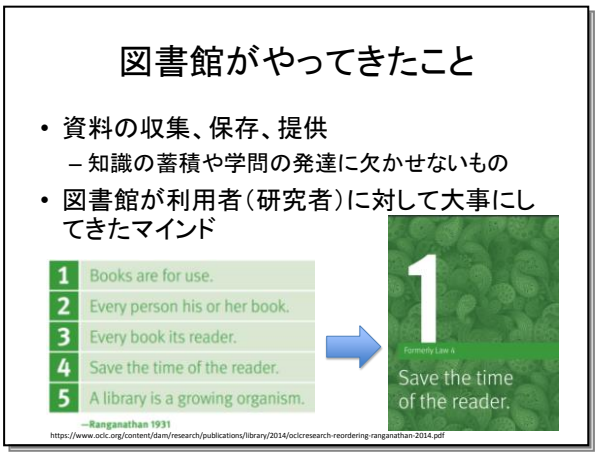
研究者情報の管理はどうなっているかというと、先ほど倉田先生が、このシステムは入力が面倒だというお話をなさいましたが、K 大学は独自のデータベースを維持していますし、さらに A 社によるシステムも利用しています。

研究評価については、評価といっても主にビブリオメトリクスを使った評価というだけの話なのですが、A 社のツールを主に使って行っています。

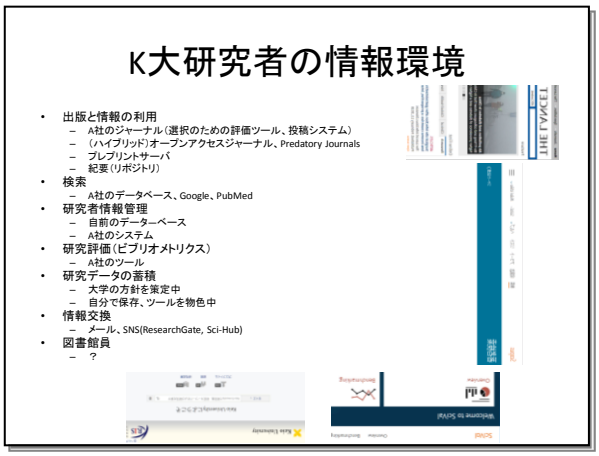
では、研究データの蓄積はどうだろうという話になると、これは大学の中で方針をつくっている最中です。先ほど倉田先生からもご自分で figshare をお使いになっているというお話がありましたが、自分で保存されたり、あるいはどのツールを使ったらいいのだろうかと考えていらっしゃったりする先生が多いと思います。

研究者同士の情報交換は、もちろん、電子メール、ResearchGate や Sci-Hub といったソーシャル・ネットワーキング・サービスで行われていると思います。

では、図書館員はといったときに、倉田先生から、インストラクションだけやっていていいのかというお話がありましたが、今のところクエスチョンということで話を進めさせていただきます。



（図 1）

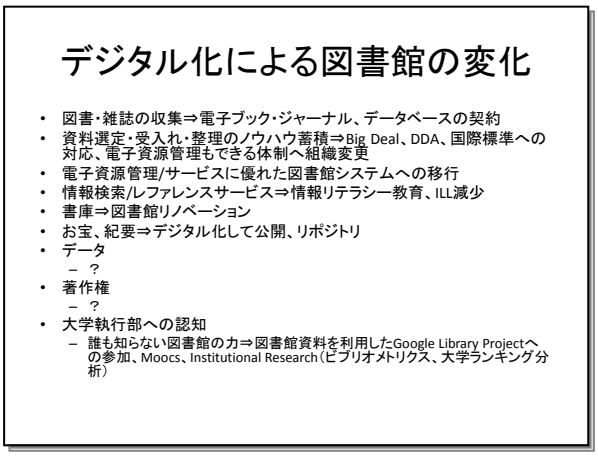


（図 2）

デジタル化による図書館の変化

先ほどは K 大学の研究者周辺の話をしましたが、今度は図書館の変化についてお話しします（図 3）。図書館は図書や雑誌の収集をしてきました。これが電子ブックやジャーナル、データベースの契約という仕事になってきました。そのため、資料の選定・受け入れや、整理のノウハウの蓄積が必要なくなるかもしれません。例えば電子ジャーナルは選択不要の big deal、本は demand driven の電子ブック、つまり出版社が提供するものあるいは、利用者が欲しいものを買えばいいというような時代になりました。つまり図書館員が持っていたノウハウというものではなく、そこにあるものを契約すればいいという時代になってきています。

図書館システムは、これらの資料を扱うために国際標準化されたシステムへと対応をしていく必要があり、電子資源の管理もできる体制へ組織を変更してきました。かなり長い時間をかけて組織の変更をしてきているのですが、図書館システムと絡めて話をしますと、現在は第 3 期のシステムに当たります。第 1 期のときはまず、それまで各キャンパスでテクニカルサービス、いわゆる発注、受け入れ、分類業務を持っていたのですが、これらの集中処理ということを行いました。第 2 期はその業務の最適化を目標とし、1998 年ぐらいから、電子資源管理システムを外付けして管理がしやすいようにしてきました。現在は第 3 期となっていて、紙と電子を統合管理できるシステムに移行してきたという歴史があります。



（図 3）

この変化の中で、図書館予算が電子資源に使われる割合が 50%を超えるようになり、紙と電子を管理するコストの増大に対応するために、図書と雑誌という区切りが元々あったところを変更して、電子資源と、いわゆる従来からの受け入れと目録という体制が本部にできました。メディアセンターは慶應義塾大学の中に六つあるのですが、本部で電子資源もいわゆる紙の資料も集中的に作業をしています。

私が図書館員として育った時代、先生方から依頼を受けて検索するいわゆる代行検索が中心的な仕事で、文献リストを出すような仕事をよくやっていました。インターネットや電子ジャーナルが普及すると、レファレンスサービスの中心はいわゆる文献の所在調査や文献検索から情報リテラシー教育へと変化しました。また、レファレンスサービスが受けていた ILL というサービスは減少しました。

書庫については、特に STM 系の図書館では雑誌、レファレンスブックスが電子化されたことによって、資料を館内の書庫に置いておく必要がなくなり、大量の資料を保存書庫へと移動しました。空いた書庫スペースを利用して、図書館リノベーションが盛んに行われています。

慶應義塾大学には、例えば『グーテンベルク聖書』というお宝があります。そういうものの他、数々の文化財があるので、それらを電子化して社会に公開しています。

データ、著作権については図書館の取り組みはまたクエスチョンという状況です。

また、図書館の変化として、大学執行部の図書館の認知が進んでいることも挙げられます。図書館の存在は大学組織の中でいつもないがしろにされてきたと私は長い間感じていましたが、例えば、先ほど述べたお宝の公開、お宝を利用した MOOCs の作製、Google Books Library Project への参加を通して、図書館の力が執行部にも見えてきた部分があります。

さらに今、Institutional Research ということで、ビブリオメトリクスや大学ランキングの分析という部分で

図書館員が力を発揮し、大学執行部の図書館の認知が進んできているように思います。私は三田メディアセンターに所属していますが、現に学術研究支援部と、塾長室という学長の補助機関のようなところのメンバーになっています。

出版社の変化

出版社の変化を見ていきたいと思います。例えば、A 社の動きがウェブサイトにまとめてあったので、図 4 に載せました。ソフトウェア開発会社の bepress、評価を行うオルトメトリック指標である Plum Analytics、実験データ管理ツール Hivebench、社会科学・人文学リポジトリである SSRN、メディアモニタリング企業の Newsflo、そして Mendeley を買収しています。研究者の研究フローに合わせた、それぞれの良さそうなものがうまく出版社に集まっているという状況があります。

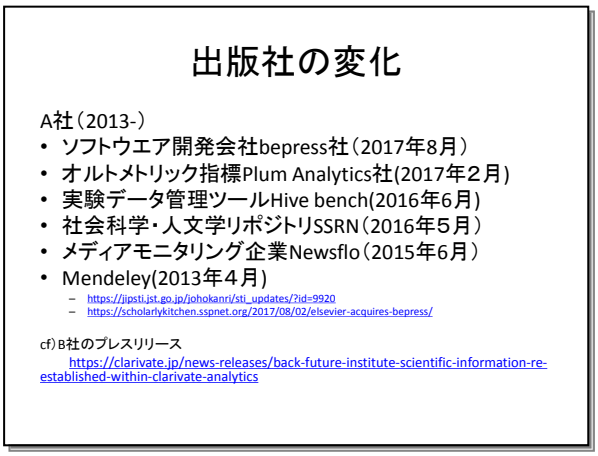
一番下にあるものは B 社のプレスリリースです。この会社も今いろいろやっている最中だと思われるのですが、今度新たにまた Institute for Scientific Information（ISI）という組織を立ち上げるというリリースです。出版社が変化してきているということが、このような情報から分かります。

私はこのような動きを、だから出版社が悪いという話がしたくて出したわけではなく、研究者もこのような出版社に依存してきたでしょうし、そこのいいところを取って、一緒にやってきたと思います。このシステムがうまく動いていなければ、科学の発展も研究者自身のキャリアもうまくつながってこなかったのではないかと思っています。出版社の方を責めるつもりではなく、スカラリーコミュニケーションの世界は適切に維持されていかなくてはいけない、それは心から強く思っています。

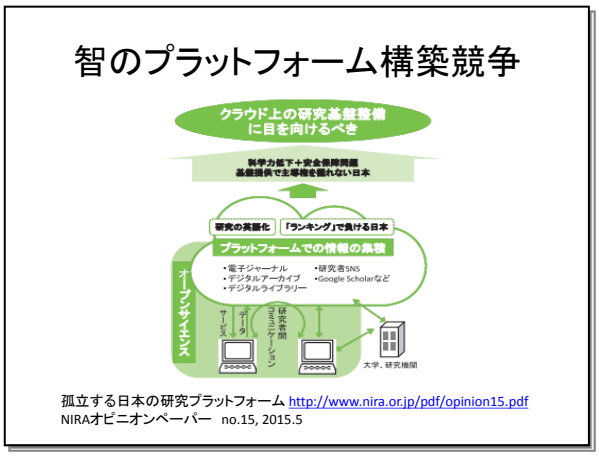
智のプラットフォーム構築競争

こういう世界が動いているということなのですが、図 5 は、2014 年から 2015 年にかけて、総合研究開発機構（NIRA）の「日本における学術智場の将来性に関する研究」に参加していたときに、最後に結論として描かれた図です。NIRA という機関自体が政策提言をするところなので、少しアグレッシブ、あおり気味な結論に持っていっているのですが、こういう世界ができればいいという絵です。

結論をそのまま引用させていただきます。「クラウドコンピューティングの発達を背景に、研究活動を支援する過程で、研究成果や評価情報を蓄積する国際的プラットフォームが形成されつつある」。これは先ほど言った、出版社が例えばいろいろなツールを持っているという話につながると思うのですが、「情報は集積することで」、今まで図書館がやってきたようなことですね、「それ自体の持つ価値を高めることができる」。持っていれば価値は高められます。しかし、「この情報の性質が、今、研究成果の情報の集積と分析の競争を引き起こしている。しかしながら、日本は英語



（図 4）



（図 5）

を共通言語とする研究プラットフォーム構築競争の中では全く主導権をとれていない。世界の研究状況が、海外プラットフォーム事業者にはリアルタイムで可視化され、日本の研究者には部分的にしか見えないという状況は、日本の科学の進展にとって重大なハンディキャップとなるだけでなく、安全保障上の問題にもなりうる。

同時に、ネットワーク化が科学のあり方そのものに大きな影響を与えつつあることにも注目すべきである。研究者をつなぐソーシャルネットワーク、市井の研究者もネットワークを介して参画しながら知を発展させるオープンサイエンス、科学的方法論の姿を変えつつあるビッグデータサイエンスなどの動きから日本は落伍してはいけない。

日本はこれまで個別の研究に注目し、大きな資金を投じてきたが、研究情報の電子化の基盤的なプラットフォーム（デジタルライブラリー、デジタルアーカイブ、電子ジャーナル）やそれを運営する人間の技能に対しては十分に注意を払ってこなかった。その状況を改めるべきである。多言語化やマルチメディア化などで、智のプラットフォームが進化する余地はまだ数多く残っており、これからでも的確な手をうつことでプレゼンスを高めることが可能である」というものです。

構築すべき情報環境を考えるために

このことを踏まえて、新たに図書館員として情報環境を考えるために、今何をしなければいけないかを少し考えてみました（図 6）。

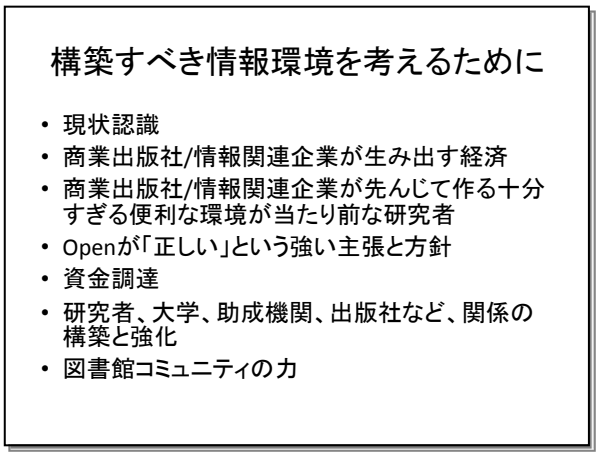
最初に、先ほど話をしたような現状認識が必要でしょう。それから、商業出版社や情報関連企業が生み出す経済をよく理解しておかなければいけません。私は「The STM Report」は出るたびに読んでいますが、英語で出版される科学技術情報の生み出すお金は 1 兆円規模だといわれています。そこで働く人たちは 11 万人いるそうです。

商業出版社や情報関連企業が先んじてつくる十分過ぎる便利な環境が、研究者としては当たり前になっていくので、これを理解しておかなくてはいけないと思います。

今日、私はお話を伺っていて、オープンが「正しい」ということを図書館員がきちんと理解していないとつくづく感じました。そういう話をきちんとしなくてはいけないと思いました。上から「オープンです」と言われるのもまた少し違うと思うのですが、それがオープンになることで、社会がどう良くなるのかということをよく考えなくてはいけないと強く思いました。資金調達は、うまく使えばうまくいくのではないかと思っています。

研究者、大学、助成機関、出版社などの関係の構築と強化も大切ですし、図書館コミュニティの力をうまく使っていければと思っています。

冒頭にランガナータンの話をしましたが、確かにデジタル化によって、利用者の行動も図書館も随分変わってきたと思います。サービスも利用者の行動も、時間の節約という方向に動いているとも話しましたが、やはり図書館の核は資料や情報だと考えています。これはランガナータンの第一の法則なのですが、資料や情報を収集するために図書館は多くの予算を使ってきていますし、今も使っています。オープンの時代になって、確かに購読だけのためにお金を全て使っていくということはもはやうまく立ちゆかなくなっていくし、それだけでは図書館の仕事を果たしていないのではないかと思えます。ですから、お金は非常に大きな力で、資料が核であるということを確信するのであれば、当



（図 6）

面オープンアクセスに向けてあらゆる支援を集中していくべきではないかと考えています。

図 7 は皆さんになじみのある絵を切り取ってきたものなので、私が何を言いたいかはお分かりになっていただけると思います。今私が関わっている大学図書館コンソーシアム連合（JUSTICE）という電子資源のコンソーシアムがありますし、オープンアクセスリポジトリ推進協会（JPCOAR）もあります。研究データ利活用協議会（RDUF）というのは武田先生が中心となって進められている研究会です。こういうものがうまくつながって、例えば、ヨーロッパの OpenAIRE のように、情報が集まるプラットフォームづくりへ力を結集して、そちらの方向へ行くのは一つの方法ではないかと思います。

David 先生のお話にあったように、オープンということによって社会が今より良くなり、また、村山先生からありましたように、図書館がデジタルな科学の中で社会基盤を築く、そういう存在になっていければと考えています。

（図 7）

第3回 SPARC Japan セミナー2017「オープンサイエンスを超えて」

質疑応答・ショートディスカッション

林　和弘　（科学技術・学術政策研究所）
Heather Joseph　（SPARC North America）
倉田　敬子　（慶應義塾大学文学部）
市古　みどり　（慶應義塾大学三田メディアセンター）

●**林**　このショートディスカッションでは、図書館の今後につながるような議論を集中的にできればと思います。何か質問や論点提示があれば積極的にお願いしたいと思います。

Heather さんにこの機会にお伺いしたかったことがあります。SPARC はもともと北米研究図書館協会（ARL）と共に始まったものですが、今は ARL から独立して活動しているという理解でよろしいでしょうか。その上で、今、SPARC と図書館の活動における関係はどうなっているのかということについて、簡単にご紹介いただけますか。

●**Joseph**　私たちは米国とカナダの ARL のプロジェクトとして生まれました。それが育って、成熟し、派生して独立組織になりました。私たちは今も図書館の会員組織です。ARL の会員の 90%が SPARC の会員にもなっていますが、ARL の会員ではないのに、SPARC の会員にはなっている図書館が他に 125 館以上もあります。ですから、会員は研究図書館以外にも広がっているということですね。アメリカでは、リベラルアーツカレッジや大学の他、2 年制のコミュニティカレッジや工科大学も会員になっていて、これらは最重要ジャーナルの文献にアクセスすることがほとんど不可能になっています。これらの機関では、オープン教科書の導入によって学生が高等教育に手が届くようになります。従って、私たちは依然としてライブラリーコミュニティと完全につながっていて、ライブラリーコミュニティのために働いていますが、対象とする図書館の幅が拡大しています。このことは、ジャーナルへのオープンアクセスを推進するのに必要な特定のアクションに向けてもっと素早く動いたり、話し合ったりする上でとても役に立っています。私たちは、研究や教育環境全体についてより包括的に見通して、図書館が重要な役割を果たせる場所を見極めていますが、その機関のタイプや、その機関に所属する人々のタイプによって異なる役割を果たしています。

●**林**　SPARC が ARL から離れたのは、ドラスチックな変化があったから、あるいはドラスチックな変化を求めたからと見る向きがあったのですが、あくまでライブラリーコミュニティとの関係性はしっかり維持した上で進められているということが伺えて、少し安心しました。ライブラリアンの方々がこの動きを気にしていたとも伺っているので、それを確認できてよかったです。

●**フロア 1**　バイオサイエンスデータベースセンターの者です。私は以前、企業でマネジャーをしていた関係でマネジメントツールに非常に関心がありました。それで今の職場に移ってからマネジメントツールを導入しようと思ったのですが、研究者の方にこういうものはあまり向かないのではないかという実感を持っていて、最近では外部の業者との間でしか使っていないのですが、先ほど倉田先生のお話を伺って、まだこれ

を使える余地があるのではないのかと思い直しているということが 1 点です。

もう 1 点、そのようなマネジメントツールや管理ツールを含めて、組織独自のものをつくるのは本当に最後の手段だということは、ここにいらっしゃる皆さんには大体ご賛同いただけると思うのです。そうなってくると、例えば世の中に出回っているクラウドを含めたツールで何を使っていくかはとても重要になると思います。研究者それぞれが独自のものを何か選んで使っていくと、どうしても互換性にも支障を来しますし、非効率だと思うのです。そういうところで図書館に、図書館だけがやるべきかは分かりませんが、補助をするというよりも、むしろ戦略性を持って「このツールを使っていきましょう」「これをまとめて契約します」といった司令塔のような役割を果たしていただくのがいいのではないかと思いました。

●**林**　ありがとうございます。大変心強い応援の言葉を頂いたように思いますが、これを受ける形でどなたかコメントなどありますか。

●**村山**　それについて、もう一度、Heather に聞きたいです。スカラリーコミュニケーション、データプリザベーション、ペーパープリザベーションの専門家として、ライブラリアン、パブリッシャー、リポジトリマネジャーがおられます。研究者でないその方々がスカラリーコミュニケーションの専門家として研究者を導いたり、コンサルタントをするという機能があるという話が先ほどの講演にあったのですが、コンサルタントにとどまらず、エデュケイトしたり、実際に訓練・トレーニングをしたり、そのようなファンクションをアメリカやカナダで見つけることはありますか。

●Joseph　同じ考えですね。ライブラリーコミュニティはデータキュレーションや研究データ管理の研修をもっと受けたいと熱望していると思います。私はジャーナル出版者としての訓練を受け、ライブラリーコミュニティに移る前に 15 年間編集者として働いていました。コミュニティのニーズが浮かび上がるにつれ、そのニーズを満たすためにスキルを新しくつくり変えていくライブラリーコミュニティの能力に、いつも私は感銘を受けています。私たちのライブラリアンや図書館の専門職員にとって、アメリカで最も急増している仕事はデータキュレーションライブラリアンや、それから派生するものです。私たちは研修を受けられるようにするためのリソースとサポートを必要としています。とりわけ、私たちは研究財団コミュニティと協議して、データ管理やデータキュレーションをサポートするのに必要なポジションに資金が出され、財団がこのようなポジションをサポートするよう、グラント要件に盛り込むよう働き掛けています。図書館にデータキュレーションライブラリアンが存在するようにしたいと思っていますし、そのために私たちと一緒に働いてくれる人なら誰とでも協力していきます。21 世紀のライブラリアンの役割は、データ管理とデータキュレーションに根差しています。

●**市古**　確かに図書館員がいろいろな場面に入っていけるチャンスはたくさんあると私も思っています。例えば、倉田先生が先ほど入力が面倒と言ったシステムについても、図書館はコミットしています。そのときに、私たちが持っている基礎知識や技術は、ID の知識、メタデータの知識など、さまざま存在します。まず学内で図書館員への信頼をつくってから、そのような基礎知識や技術を図書館員は使えるのだよというところをうまく見せていけば、チャンスはまだたくさんあると私は個人的には感じています。

●**林**　これに関連しても構いませんし、他の論点、いかがでしょうか。ICSU-WDS の方、お願いします。

●**フロア 2**　私もやっているのですが、ドメインの研究者がそれぞれのデータセンターを持っているという大学は、京都大学や東京大学など、たくさんあると思

うのです。大学のドメインのデータセンターと、図書館におけるデータのアクティビティはどういう関係になるのでしょうか。

●林　それはデータサイエンティスト、データライブラリアン、データキュレーターとは何ぞや、そのスキルセットは何かというディスカッションになります。それを明快にご説明いただける方は倉田先生でしょうか。

●倉田　スキルセットというよりは、大学内でどのようにデータを管理したいかということになると思います。例えば、理工学部などでは、図書館とは全く別にデータセンターをつくった方がいいと考える大学ももちろんあると思います。では、図書館がそこに本当に加われるのかは、そのデータセンターを大学の中でどのように位置付けるかという問題だと思います。

それを言うのでしたら、高エネルギー物理学分野など、分野のデータセンター自体が大規模に構築されている分野や領域は幾つかあるわけで、そういう方向に進むこともできるわけです。村山先生のところもそうです。つまり、特定の分野ドリブンでデータセンターをつくることもできるし、大学内でつくることもできるし、図書館がそれを主導することも不可能ではないというところで、いろいろなレベルでいろいろな運動が起きてしまうのだと思います。

それをどう統合するのか、調整するのかというのが、今、見えていないというところが問題だと思うのです。そのためには全体のランドスケープ、全体をどうしていきたいかという目標をもう少しみんなで共有できた方がいいのではないかと私は思っています。

●フロア 2　ありがとうございます。去年、京都でWorld Data System（WDS）関係のシンポジウムをしたときに、イリノイ大学の図書館の方を招待したのです。そのときにおっしゃったことは、大学の中にあるデータセンターは確かに専門性があって、利用者のニーズに応えたきめ細かいソリューションを提供できるけれど、一番の問題点はファイナンシャルだということです。研究者は研究第一なので、そういうデータサービスにはなかなかお金が来ない、これからもどんどんそうなることが予想される、そういうときに大学全体として、図書館とドメインのデータセンターが一緒になってファンディングなどのサポート体制というか、永続性のあるデータサービスをエスタブリッシュするシステムを考えるべきではないかと思うのです。

●倉田　そのとおりだと思います。今、それが完全には見えていないのですが、個々の図書館も国立図書館も含めて、今後ものすごい量のデータになっていくのは明らかで、一体それをどういう形でみんなで作り上げていけるのか、答えは一つでなくてもいいと思うのです。段階的にでもいいと思います。でも、あまりめちゃくちゃな方向に行かないように、できるだけ段階的にでも目標に向かっていくためには、もう少し最終的な目標像が共有できた方がいいのではないかと考えています。

●林　それがまさに Paul さんの言っているオープンサイエンスのシステムというものですね。それが共有できてからでなければビジネスも回らないという話です。Heather、どうぞ。

●Joseph　図書館とデータセンターが協力するのは大事なことです。日本の大学にコロラリーがあるか分かりませんが、アメリカでは、委託研究のオフィスと研究オフィスがあります。図書館と委託研究のオフィスは、特にデータとデータ管理の考えを巡って、異なる種類の協働関係を築く必要があります。現時点では、非常に弱い協力しかありません。私たちは問題を詳しく調査し、必要なリソースについてもっとよく理解するために、より密接なチャンネルを築く必要があります。他に関与してもらうことを考えるべき関係者は、学術団体です。私たちのデータライブラリアンが自分

に足りないと感じることがあるスキルセットは、データやその特性についての学問領域特有の知識です。学術団体は、学問領域特有の専門知識をもたらす上で、新たな種類の役割を果たせるでしょう。従って、私たちはその領域における従来からの重要な貢献者である関係者と共に、従来のものとは若干異なる新しい種類の協力を模索しなければなりません。

●**村山**　その意味で、深貝先生に大学図書館長の立場からの意見を伺いたいです。データライブラリアンの仕事は日本に今ないのですが、大学の学部ファカルティでつくったデータのメタデータなど、ある程度ディシプリンから離れた学術情報管理として、データに関わるようなワークフローなどを日本の大学図書館でつくることはハードルがかなり高いでしょうか。実験的なことやパイロットプロジェクトをするという道はないものなのでしょうか。

●**深貝**　まず、私はもはや図書館長ではありません（笑）。

●**林**　だから言えることがたくさんあると思います。

●**深貝**　後でパネルのときに少しお話しさせていただきますが、今の話の限りで言うと、大学図書館でデータの管理に関わることまでお任せするというのは、今の大学図書館の方々が置かれたさまざまな用務とスタッフの配置からして難しいです。それは大学図書館がやるかやらないかではなくて、大学の管理運営の責任を持っているセクター、つまり管理職の方々が、自分の大学の機能をどこで発揮すべきかと考えたときに、学内の特色あるデータを世の学術コミュニティに積極的に出す必要があると判断するか、どう取り込むかという問題に恐らくなると思います。

　従って、研究に重点を置いている総合大学であれば可能かもしれませんが、研究に重点を必ずしも置かないような大学であれば、学部が専門性があってユニークな学部に絞られているような大学であれば考え得るけれども、そうではないところだと相当難しい話だと思います。

●**村山**　結局は、「論文アウトプットだけではなく、データも本学の重要な学術情報資産であるから図書館として取り組むべきだ」という、大学のエグゼクティブ、ヘッドクオーターの意思決定がないと恐らくは動かないと想像しているのですが、そういう流れをつくる・・・すみません、自分で言っていて、そういうこともハードルが高いですよねと結論が出てしまいます（笑）。

●**林**　ぜひまた Paul さんと Heather さんにお越しいただいて、大学の経営者向けに、科学とは何ぞやから入ってこの話をしていただきたいという流れになってきたと思います。それでは後半のパネルディスカッションにそろそろ移りたいと思います。どうもありがとうございました。

第3回 SPARC Japan セミナー2017「オープンサイエンスを超えて」

パネルディスカッション

深貝　保則　（横浜国立大学大学院国際社会科学研究院）
Paul A. David　（Stanford University）
村山　泰啓　（情報通信研究機構/ICSU-World Data System）
Heather Joseph　（SPARC North America）
倉田　敬子　（慶應義塾大学文学部）
市古　みどり　（慶應義塾大学三田メディアセンター）

●**林**　それでは最後のパネルディスカッションに入りたいと思います。パネルディスカッションはモデレーターを元横浜国立大学図書館長で SPARC Japan 運営委員会委員の深貝先生にお願いしております。まず深貝先生からの問題提起を皆さんに聞いていただく時間を設けています。その上でパネラーにご登壇いただき、パネルディスカッションにしたいと思います。まずは深貝先生、よろしくお願いします。

討論への問題提起：オープンサイエンスの知識論を考える

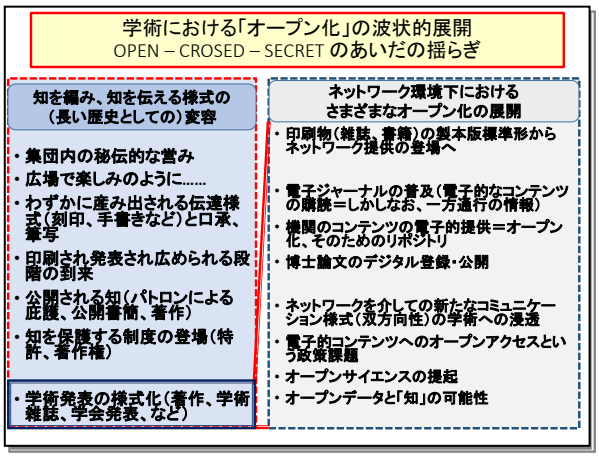
●**深貝**　午前中の Paul David 先生は歴史的な話とともに、今、起きつつあることは何なのかを、村山さん以降、Joseph さん、倉田先生、市古さんは、データベースとして何が整えられ、さまざまな図書館の機能が何をするのかを話されました。これらをどう結び付けて考えるのかという形で、大づかみの話をしたいと思います。

学術における「オープン化」の波状的展開

最初に、学術におけるオープン化が進んでいますが、出発点はある意味ではオープンでした（図 1）。古典古代、ソクラテスは広場で討論していました。そういうところにおいては極めてオープンに知識というものを語り合っていたわけです。ただし、ソクラテスよりほんの少し前のピタゴラス教団は秘密主義で、内輪だけの知識でした。それに対し、知識を表に出して知を楽しむ、オープン化はそこから始まっていくのです。

そのようなものが少しずつ書きためられるわけですが、近代の近くになるまで、中世の終わりぐらいまでは、書かれた知識というものは大変でした。中世では、手書きの文書が訓練を受けた手書きで書き写されるから本のようになるというものでした。そのようなものが、近代における技術革新の結果、やがて広まっていきました。

大量生産される書物になってくると、誰が書いた知識であり、誰が書いた知識を誰が出版していったのか、知識の持ち主は誰で、知識を出版した人々の努力はどう報われるのかという話になりました。従って、著作



（図 1）

権や特許、その他いろいろな権利関係の問題が整えられ、法的にも支えられるということになってきました。

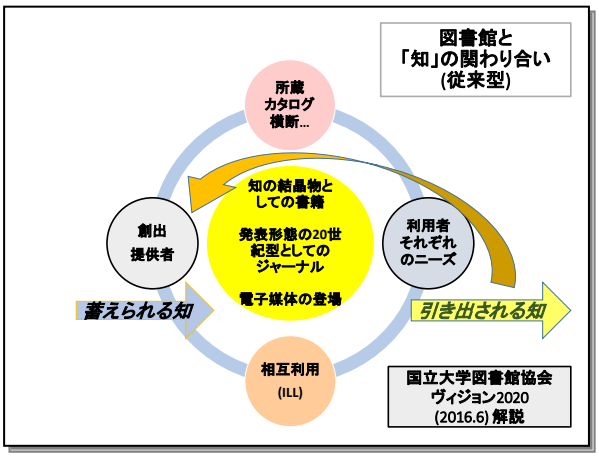
そのようなことが進み、19 世紀から、徐々に学術が細分化し、制度化していきます。20 世紀に入り、今日に似たさまざまな分野がはっきり出来上がってきました。ヨーロッパにおいては近代の初めくらいまでは、学部といっても非常に大づかみな学部でした。細かく学術が分かれているわけではありませんでした。従って、特殊なものを除いては、専門分野の学術雑誌などというものはなかったのです。包括性の高い学術であったのですが、19 世紀ぐらいになるとだんだんそれが細かくなり、近代化していく中の日本でも、現代の学部に至る、細かく特徴を持って目的を限定した学術がどんどん進んでいきました。

その中で 20 世紀に入ってから学術雑誌が徐々に中心的機能を果たし、20 世紀終盤になって、それが電子化され、電子ジャーナルを軸としながら学術が動くという姿が今、展開しているのです。

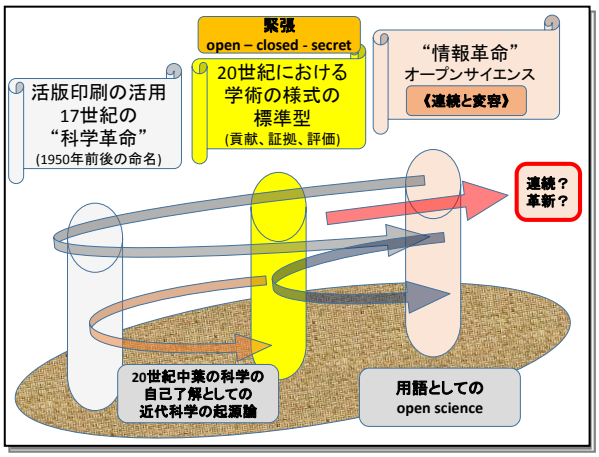
ネットワーク環境が進む中でオープンを語るわけですが、そのように考えると、印刷技術の飛躍的発展は大きな革新であったはずです。今日の午前中、David 先生もその点を踏まえてお話しくださいました。そのようなものから続いて電子ジャーナルへ、そして電子的コンテンツをジャーナルに収録されたものだけではなく、それ以外のものも載せていこうという形でリポジトリが機能していきます。

日本では特に図書館をベースとしながら、21 世紀に入った当初からオープンアクセス化ということで、さまざまなコンテンツを入れようということになり、しかも日本の場合は、研究機関あるいは学術教育機関はできる限り自前の仕組みを持とうということで、機関リポジトリが非常に多く整備されています。機関リポジトリが先に出来上がっていった上で、その後でオープンサイエンスという課題が突如現れたように見えます。今日の午前中、村山さんからもそのような話がありました。国内の普通の感覚からすると、2010 年代半ば近くになってから急にオープンサイエンスが飛び出してきた。有川節夫先生が内閣府の委員会の座長であって、有川先生のバックグラウンドからして、図書館にまず頼ろうとなったので、大学の管理部門の人々がよくオープンサイエンスを理解しないうちに図書館の仕事にいつの間にかなった。そして、図書館のスタッフに、学内で説明するときにご苦労いただいた時期が一時期あります。しかし、その次の段階に今ようやく来つつあるわけです。

オープンサイエンスの課題やその他のことを受けて、2016 年の国立大学図書館協会の総会で「ビジョン 2020」を決めています。その「ビジョン 2020」の設定を私は裏方として手伝わせていただいて、そのときの少し前からつくったものが図 2 です。従来型の図書館は、書物や雑誌に盛り込まれている学術的知識を必要に応じて引き出して、そこで考えた上で、考えた結果が新たに付け加わった上で書物や雑誌になり、図書館に収めるのにふさわしいものは収蔵される、こういう



(図 2)



(図 3)

サイクルです。一つの研究機関で足りなければよそから借りるという話でした。

そこで次の話です。これは詳しく話します。少なくとも今、現役の研究者にとっては、1980 年代にはやった学術スタイルが中心だと思っている人はもはやロートルです（図 3)。1990 年代ぐらいになってからようやく整ったような電子ジャーナルを使って雑誌論文を読むことになじんでいる世代が、院生を指導するような、現在の中心となる研究者です。自然系の人でも、それより少し前は電子ジャーナルはほとんどなく、雑誌というのはコピーするものでした。

電子ジャーナルの展開に伴って、評価システムとしてのインパクトファクターというものが広まりました。そうなると、電子ジャーナルにおけるコントリビューションとエビデンスをどう揃えるか。なお、これを巡っては、研究不正あるいは研究公正についていろいろな問題が伴っています。

ウェブ上で情報が得られるという技術的な面の限りは、電子ジャーナルからオープンサイエンスへの連続性があります。しかし連続性があるけれども違うかもしれない。違うかもしれないというのは、自分の研究成果を論文に書く、これは自分が一番乗りになることがいいことだと普通は思ってやっているわけです。一番乗りになろうとしているときに、自分の途中プロセスのデータをオープンにすることに、かなりの研究者は不安に駆られます。そうすると結局、連続であるようでありながら違ったことになるかもしれないという戸惑いの中にいるので、学術は今、連続的か不連続か、変わったことが起きつつあるのかという状況に置かれているのです。

過去を振り返り、現代の知識を論じる

この意味を考える際に、グーテンベルク活版印刷、そして 17 世紀の科学革命と呼ばれる時期の様相が大いに参考になります。20 世紀、21 世紀にかけての、情報の伝わり方、情報に関するコミュニケーションの仕方のラジカルな変化と似たことが 17 世紀に起きていたと擬似的に考えて、そのときに起きたことは何をもたらしたのかを考えることが一つの手掛かりになるのです。

20 世紀半ばに第 2 次世界大戦が激しくなったとき、あるいはそれが終わったときに、世界体制をどうするかという国家的スキームの中に科学者が引き込まれていった時期があり、そのときに学問的知識とは何かということを真剣に考えようとした人々は、時に「科学革命」と呼んで、17 世紀を論じはじめたのです。

“scientific revolution”という言葉は 17 世紀にもちろんあったわけではなく、20 世紀半ばに学術用語として登場しました。20 世紀半ばに 17 世紀を見ることによって、自分たちがやっている学術とはどういうものであり、知識とはどういうことかを考える手掛かりにしようと考えた人々が、わずかですがいたわけです。

そこで、現代の情報革命の下でのオープンサイエンスの課題についてですが、いわば過去にブーメラン的に二重に飛ばしてみて、そして今は何だと考えてみることは手掛かりになるかもしれません。David 先生のお話の半分方はこういうストラテジーの下にあると思います。20 世紀半ばからの過去へのブーメラン、それがあることを知った上で、21 世紀初めの段階で 17 世紀にブーメランを投げ、さらに 20 世紀半ば以降の現代につながる科学に対して、少しばかり距離の近い過去へさかのぼるブーメランを投げてみる。その中で現代は何が起きているかを見てみようというわけです。

ブーメランを後ろにばかり、過去にばかり投げていても何ですから、先に進んで考えて、将来に向かってブーメランを投げてみると、戻ってくるのか、それとも飛び立っていくのかよく分かりませんが、過去のことを参考に、これから起きつつあるオープンサイエンスの中から得られる知識は一体何かを考えてみよう、今やっているのはそういうことだと思います。

そのような次第で、問題提起としては、オープンサイエンスの「知識」を考えるのではなく、オープンサイエンスの「知識論」を考える、知識のあり方そのものをどう論じるかを考えるという形でフレームワーク

を出させていただいています。

“open science”という言葉は、私が気付いた限りでは、18 世紀に 1,000 ページの本に一回だけ出ていましたが、現代のオープンサイエンスとは全く違った意味だと思われます。また、19 世紀末の段階でケンブリッジやオックスフォードにオープンサイエンス・スカラーシップという奨学金があったようですが、これは何だかよく分かりません。

学術論文としてオープンサイエンスを表題に掲げたものとしては 1980 年代半ばに、オープンサイエンスとクローズドサイエンスという対句で出ています。その後に David 先生が、“The Economics of Open Science”のように、引用符付のオープンサイエンスがタイトルに付いた論文を何本か、1990 年代から既にお書きになっているので、われわれはオープンサイエンスとは何かを語る、そのハードな議論をしている先生をお招きして今日は議論できているわけです。

なお、こぼれ話として、18 世紀にもリポジトリがありました（図 4)。ただし、印刷体です。

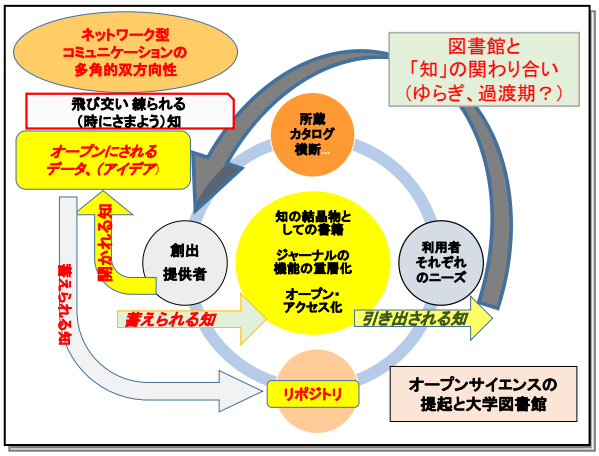
図書館は今、新しい課題に向かいつつあります（図 5)。本当はまだ 1 枚スライドが残っているのですが、それについて話す時間がありません。モデレーターは少しずるいことができるということで、そのスライドをこれから小出しにして、パネリストたちにお考えをお訊ねするという形で進めたいと思います。

●林　深貝先生、どうもありがとうございました。それではパネルディスカッションに移りたいと思いますので、スピーカーの方々は前の方にお越しください。

ここからはマイクをモデレーターとしての深貝先生にお渡ししたいと思います。

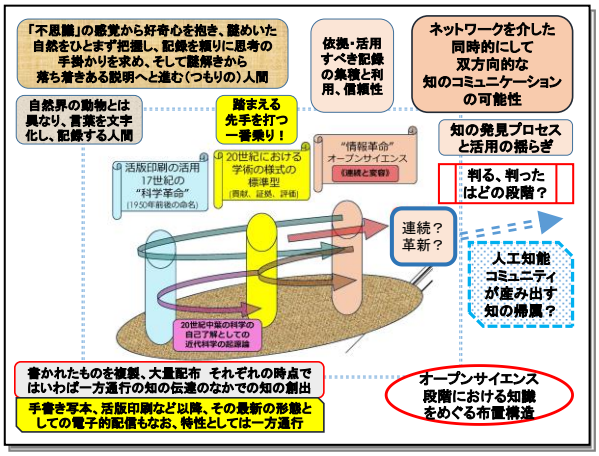
●深貝　では、少しずつスライドを出しながら、いささか怪しげな話をします。私は思想史研究者で、社会倫理学なども研究しています。倉田先生は最近のメディアをふんだんに使っておられますが、私の方はほとんど使っていません。最初の何年間かの論文は手書きで、本当にのりとはさみでカット＆ペーストをした世代なので、まだ新しい方式になじんでいるわけではありません。

さて、自然界における動物と人間とはどちらが優れているかというと、普通、人間が優れていると思われ



（図 5）

（図 4）



（図 6）

ています。動物の方が優れていることはいくらでもあるのですが、人間の特徴の一つは記録を残すことです（図 6）。記録を残して、自分の記憶だけではなく他人の記憶まで借りてものを考えるという、いわば特権的な仕事を行っているわけです。

中でも、科学的な考察を人間は行っています。「不思議だ」という感覚から好奇心が芽生えて科学になります。人間が「けしからん」と思ったときには科学になりそうにありませんし、「素晴らしい」というものでも科学になりそうにありません。ところが、「不思議だ」と思って、不思議さをきちんと詰めていくと科学になります。その場合に、ただ「不思議」では訳が分かりませんから、試してみる、実験する、あるいは記録を手掛かりに誰か同じようなことを考えた人の考えを参考にする、さらには議論をする、不思議で会話を楽しむということをやっているわけです。

というのが今までなのですが、「不思議だ」という感覚から「判りたい」という場合に、今、大量に集められようとしているデータはどのように役に立つのでしょうか。あまりデータが多過ぎると、必要なデータはどうやって取り出せばいいのでしょうか。

これは村山さんか Joseph さんか、どうでしょうか。

●村山　科学による知識の生成サイクルにおいて、データがどういう位置付けかという問題を新しいものとして捉えるかどうかだと思うのです。データというのは特別なものではなく、何かの観察や実験をしたときの対象や実験結果、あるいはそれを蓄積していった情報資産ですから、例えば昔のケルビンなどが運河を流れてくる潮津波などを見て流体力学の発想のアイデアを持ったというのも、目で見た情報を直接自分の脳が感知したデータだと思います。それをセンサーで捉えるのが今の、モダンなやり方です。

最近の科学においては、「不思議」の感覚と科学が分離しているのではないかというのが私の意見です。つまり、空が青いという生活感覚が、太陽光のスペクトルのうち赤い光のところが吸収されて青いという科学的知識に結び付くときに、空の撮影写真や光のスペクトルはデータとして有用なのですが、これを宇宙空間においてプラズマが吸収する周波数帯と言うともう生活感覚ではありません。つまり、皮膚感覚の不思議によって科学がドライブされているサブジェクトではなくなってしまって、よく言われるたこつぼ、狭い領域の中での専門知識のやりとりの中でだけ発生する疑問、これを解決することが科学的なターゲットになることが多くなっているように思います。そこにおいて専門的なデータは非常に重要な価値を持つので、「不思議」という皮膚感覚から遠ざかった科学の営みを、今われわれはやらざるを得なくなっています。

それを否定すべきかというと、そうではありません。それは進化した科学であり、かつ、私は政府に、公的な資金を必要な科学に投資しなければいけないのではないかと問い掛けている立場なので、そこにおいて再構成された科学が何を生み出せるのか、その中に科学者自身の興味や不思議を見つけながら、人類にとってプロダクティブな営みが次にどうできるのか、そこが新しいモードになってきているのではないかというと、その可能性があるのではないか。そのための知識資産としてのデータをわれわれは位置付け直して蓄える必要があるのではないかというところまでは何となくイメージができています。

●深貝　先ほどあのような質問をしましたが、不思議が好奇心で科学につながるというのは、実は特殊な想定であるというのは別のベクトルからも言えるわけです。旧い時代で言えば、不思議の感覚は星占い師や魔女やシャーマンの世界で、人をうまく説明して丸め込むという話になるわけですから、実は不思議が科学になるということはなかなか大変な想定があるわけです。科学というものが目に見えてあるからこそ、不思議を好奇心で、科学で説明してみようと思うわけです。

今の村山さんの話からすると、先端の科学者にとっては不思議はもはや切り離された世界ですが、子どもがやがて科学の道に進みたいと思うときには入り口ぐ

らいにはなるのではないでしょうか。

●**村山**　本筋の話ではないかもしれませんが、科学というのは、人類が持っている、ある証拠とロジックの集め方に基づいた一番もっともらしい結論の導き方で、かつては、仏教など何らかの宗教行為が、国や集団の統治原理だったのが、今は科学が統治原理として機能しています。そこにおいて一番もっともらしい将来の羅針盤、将来指針をつくっていくための論理根拠が科学しかないだろうということです。科学というのは人類の重要な行動原理として育てなければいけないし、キープしなければいけないということを、為政者というか、政策決定のコミュニティで理解していただくことが人類のために非常に重要だと思います。

●**深貝**　どうもありがとうございます。

●David　私からはこの議論に、時間という次元の話を付け加えたいと思います。私たちは、自然界のデータについても、そして実験によって生成されるデータについても、多くのデータを集めたいと思っています。自然界のデータの場合、大事なのは、そのデータが今必要なのかどうかという点です。自然界について研究している皆さんは、すなわち反復する現象について研究しているわけです。その現象は消滅しません。そして、もし消滅する現象について研究するのであれば、その現象を観察結果として記録する際には優先順位を検討する必要が出てきます。自然界についての観察結果にアクセスできるまでの時間というのは、私たちが観察結果を理解しようとしたときに、何が私たちの観察能力を混乱させたり、妨げたりするのかを予測することによって変わってくるでしょう。つまりこの先、火山噴火のような自然過程が起こり、成層圏や対流圏について今と同じような観察を行うことは不可能になるだろうというのであれば、将来使うためのデータを今記録することもできます。あるいは、火山の噴火が次にいつ起きるかを予測する場合など、一定期間内にそのデータを効果的に使えるだろうというような場合も、今すぐデータを取る十分な理由になるでしょう。そして、次の噴火が起こるまでに、その特定の現象についての観察を、より集中的にやっていこうと考えることになると思います。

同様の例として、例えば何らかの種が絶滅しかけていて、その種の習性や行動、身体的特性が観察できなくなってしまう。しかし、その種の由来となった遺伝子は遺伝子プールの中にあり、他の遺伝子との相互作用があるだろうということが分かっていれば、その種が絶滅する前に観察したいと考えるでしょう。

正直なところ、私は経済学者ですから、常にリソースの配分を考えています。観察能力が乏しいとなると、観察にどんな優先順位を付ければいいのでしょうか。一つ基準となるのが、今後その観察結果を使って何かしようと思ったときに、それが観察できる状態にあるだろうかという点です。逆に、その観察対象がこの先消えずに残るかどうかといった点についても考えなくてはいけません。このように、今、取らなければならないデータはどれなのかという検討は非常に複雑なプロセスです。また、これは社会現象についても同じことが言えます。そういったことを考えておく必要があるのです。

●**村山**　一部の経済学者や社会科学者が、将来の市場の外部性は時を超越しているということを論じていたことを思い出しました。同じような研究結果を私たちも見いだす必要があるのかもしれません。

●David　ここで深貝教授が先ほど言われた「オープンサイエンス・スカラーシップ」という言葉について補足したいと思います。「オープンサイエンス・スカラーシップ」というのは「オープン・スカラーシップ」という言葉を変形させたものなのです。オープン・スカラーシップというのは、誰でも応募できる奨学金のことです。エセックス州出身である必要はありませんし、プロテスタントの牧師の子どもであるといった必

要もありません。ただ、それがサイエンスの分野におけるオープンな奨学金であるという意味で、オープンサイエンス・スカラーシップといわれるようになったのです。ですから、サイエンスの奨学金であり、オープンな奨学金であるというのが正しい理解なのだと思います。イギリスで生まれ育ってはいないけれど、そこに長く暮らしてきた私にとって、その言葉が実際に何を意味しているのか、そして、過去にはどういう意味で使われていたのかというのを知らないことも多く、発見の連続です。

●**深貝**　林さんから午後のお 3 人に関しては質問をしていただきましたので、非常に簡単にですが―。日本は機関リポジトリを含めていろいろ持っています。例えば慶應義塾大学における、グーテンベルクの活版印刷の聖書など貴重なものを持っています。慶應では立派にそれを解析し、オープンにし、大きなプロジェクトを動かしていらっしゃるわけですが、国内には実はそうではなく、眠ったお宝を放ったらかしにしているという状態があります。これは、図書館業界に限らず、コレクションを持っている研究機関の宿題なはずです。ただ、その場合に、宝があるからといって、日本語でカタログを作ってもうまく伝わるかどうかは定かではないという問題があります。

Joseph さんに教えていただきたいのですが―。日本の研究者が英語で書いた論文は、国際的にもチャンスがあれば読まれるわけですが、そうではない日本の学術情報はどの程度、英語圏で、特にアメリカで知られているのでしょうか。学術情報に関わる人たちの間で、日本の学術情報の伝わり方はどの程度の出来栄えだと言われているのでしょうか。

●**Joseph**　それが英語という言語で書かれていなければ、アメリカでも、そして国際的にも、その学識はしかるべき評価を得られないということが多いでしょう。オープンであることの有用性の一つ、そしてオープンリサーチの利点となり得る点の一つが、その素材を翻訳できるというところです。オープンな素材であれば、翻訳に制限がかかることはありません。そのため、少なくとも理論上は、学術論文や学位論文が、元々は日本語で書かれたものであっても、真のオープンアクセスの状態であれば入手できることになり、英語に翻訳された形を含めて、より幅広い読者にリーチする可能性が高くなります。こうしたサービスは提供可能だと思います。しかし多くの場合、オープンアクセスを推奨すると、何かを無料で手に入れたいだけだと言われてしまうのです。けれども、翻訳は非常に価値のあるサービスとなる可能性を秘めています。積み重ねられた重要な研究に対して、誰かがそのサービスを提供してくれれば、その知識は、それまで届いていなかった、より多くの、しかるべき読者にリーチすることができるでしょう。

●**深貝**　今の点に、倉田先生、市古さん、何か付け加えられることはありますか。

●**市古**　他の方にお伺いしたことがあるのですが、シニアな研究者は日本語で書かれていても読むことができて、しかも日本語でも書けます。しかし、やはり研究過程の人たちは日本語を読むけれども日本語で発表することはありません。ですから、逆に言うと、日本の発表媒体で発表することはないのですかとお伺いしたところ、それはよほどシニアになってこないと、そこのハードルは高くてできないという話は聞いたことがあります。

言語の問題というのは、今、何となく競争の時代でいつも大学の中で話題になるのですが、今オープンであれば翻訳ができるというのは目から鱗な感じがしました。先ほどいろいろな著作権のマークが出てきましたが、やはりオープンの価値というのは、そういう面からもうまく広げていきたいなと思いました。

●**Joseph**　今日は歴史学者がいらっしゃるのでお聞きしたいのですが、私が翻訳についてよく耳にする皮

肉な点の一つが、歴史学の分野には特に、翻訳は二次的著作物であるのでオープンアクセスを認めてはならないという規律があるということです。そこで、先生はそれを良いことだと思ってらっしゃるのか、そうでないのか、お考えをお聞かせ願えないでしょうか。

●David　翻訳は二次的著作物です。というのも、優れた翻訳には、原文の思考を表現するための解釈が伴うからです。ニュアンスを表現するには解釈的な翻訳が必要です。詩を翻訳することの難しさについても、非常に数多くの文献があります。詩を翻訳することは不可能なのです。詩を読んで、そのリズムや音色、感情などを伝えようとすることはできますが、それは創造的な行為です。単純な俳句の一首であっても翻訳しようと試みたことのある人ならば、これに同意してくれるでしょう。

オープンなクリエイティブ・コモンズ・ライセンスの問題は、これは自分の作品を他の人が使って二次的な形式で複製することを制限する意図でつくられたシステムなのに、作品の一部を省いたり、文章を操作したりすることで、何か別のものを創作することができるということです。私が先ほどお見せしたスライドにあったのは、クリエイティブ・コモンズの中でも、商業的利用や二次的著作物を認めない、帰属のみのライセンスです。それ以外の場合は、他の人があなたの作品やあなたの名前を使って、何かを言っているように見せることができてしまいます。ですから、もしかしたら私たちは、クリエイティブ・コモンズ団体に言って、翻訳とそれ以外の二次的著作物を区別しないでほしいと頼むべきなのかもしれません。

私自身は、私の著作を翻訳したいと言われれば承諾していますが、完成した翻訳を私に送るようにと求めています。そして、その言語のネイティブスピーカーを自費で手配して、私が書いた元の文章を理解した上で読んでもらうのです。もしその翻訳チェックをしてくれた人が、翻訳に間違いがあると伝えてきたら、私はジャーナルに連絡して、翻訳は使ってもいいが、ここは変更してほしいと伝え、変更箇所のリストを渡します。最初に提出された版の翻訳を使用することは許可しません。こんなふうに、著者が翻訳をコントロールすることも可能なのです。これが自動化できるかどうかに関しては、私は疑問を持っています。しかし翻訳と、法律におけるその位置付けというのは、興味深い、重要なポイントだと思います。

●**深貝**　翻訳と解釈は話をしだすと長くなってしまいます。遣唐使のころからの漢字文化の翻訳と明治維新以降の西洋文化の翻訳で、日本が何の知識をつくったのか、文明をどのように培ったのかというような話になってしまいまして、ちょうど慶應の福沢諭吉以来のトピックスですが、それはそういう話になるということだけを言っておきます。

話のきっかけになると思いますので、スライドでまだお話ししていなかったことを紹介します。書かれたものは、グーテンベルク活版印刷から電子ジャーナルに至るまでは一方通行でした。基本的には一方通行で、出されたものを読んだ上で、今まで書かれたものを踏まえた上で考える。先に手を挙げると予算を取りやすいですから、先に研究プロジェクトを立てて、そして結果が一番乗りというのが意味ある学術上のコントリビューションだということになっていた面があります。もちろん、関心のあり方の中で何を解明すべきかというのが、それぞれの学術コミュニティの中での研究課題の大きな流れとなってきたわけです。

ところが今、ある意味では大きく変わりつつあるのは、オープンサイエンス、つまりオープンな討論が可能だというコミュニケーションのあり方を巡ってです。日常のコミュニケーションでも手書きの手紙を書く人はどんどん減りはじめています。e メールをはじめ、もっとカジュアルなやりとりを行っています。学術はそんなにカジュアルではやはり厳密性に欠けるので、もっと別のスタイルに当然なるはずでしょうけれども、それでも双方向のコミュニケーションが行われるようになってきています。地理的に離れた人があたかもフ

ェイス・ツー・フェイスで対話をし、その結果として、会ったこともない人同士が何千人も集まった共同研究論文も出はじめているわけです。

そういう中での知の可能性についてですが、しっかりした研究機関やしっかりした学術団体が提供しているデータだったら本当に信頼できるかというと、本当はよく分からない面もある。意図的不正がないとは言えませんし、厳密なものであっても特殊な条件であるデータであって、その限り有効なデータを、特殊な条件でなくても通用するデータだと錯覚して使うと変な結論が導かれるなど、いろいろなことがあるはずです。

そのようなわけで、知識として分かった、分かっていないの判断基準は一体何なのか。ある意味でチャンスが広がっているが、ここまで来たら分かったということが分からなくなりつつあるようでもある。そういう意味で、知識というものが何であって、学術をどう進めるのか。オープンサイエンス段階でオープンな討論をすることは、クローズドで隠し合っているよりはましということは明らかに言えそうですが、その上でどこまでやったら分かったことになるのかという問題です。

今までの学術に比べて、これからのそれは連続的なのか不連続なのか―。日本では将棋や囲碁などで人工知能が生み出した新しい手がプロの棋士を打ち負かしはじめており、そうすると人工知能に身体があるのか、主体性があるのかはよく分かりませんが、それは別としても、人工知能が生み出す結果的な知識というものはあるわけで、人工知能のコミュニティの中で訓練され、結論が得られつつあります。

しかし、人工知能の一つ一つには、開発者あるいはスポンサーがいます。そうすると、人工知能コミュニティが生み出した知識は、人間の側の個人もしくはさまざまな運営団体、主体の側の一部の所属物であっていいのか、そうではないのか。人工知能は、人間がつくり出した擬似的な子どもであって、子どもコミュニティのような人工知能コミュニティであるとすると、子の親離れが始まったのちには、親がコントロールすべきではない独自の世界かもしれません。それとも、人工知能コミュニティはやはり人間の外付けハードディスクのようなコミュニティで、本来の脳は人間の脳だから、人工知能コミュニティが生み出した知識は人間の持ち物だと言うべきでしょうか。どちらでしょうか。

武田先生は人工知能学会の倫理コード制定のブレーンですから、先生と昨日少し議論したのですが、こういうややこしい問題に突入すると、人間の開発した知識は、どちらがジャーナルに先に載ったかというようなことを超えた問題にそろそろ出合うかもしれないということだけ、おまけとしてお話ししておきます。今日これを討論してくださいということではありません。

さて、あと 30 分ほど時間がありますので、皆さんから質問・コメントなどいかがでしょうか。

●**フロア 1**　JST の者です。モデレーターのトーンに合わせて哲学的な質問をしたいと思います。実務者としてオープンサイエンスを見るときに、これは物事が変わるのだという認識をしがちなのですが、なぜサイエンスがクローズからオープンに変わっているかというと、これは単にオープンサイエンスを可能にする技術革新があったからにすぎないという認識も成立します。そうすると、もし 17 世紀に現在のようなインフォメーション・コミュニケーション・テクノロジーがあったならば、17 世紀からオープンサイエンスをやっていたのだろうかという質問も出てくるわけです。

そこで、パネリストに個人的な意見をお伺いしたいのですが、オープンサイエンスというのはそもそもの科学のあり方にだんだん近づいていくということを意味しているにすぎないのでしょうか。それとも、技術革新に伴ってサイエンス自体が本質的に変わろうとしているのか、どちらでしょうか。

●**深貝**　これはパネリストのどなたかということで、指名なしですが、では David 先生。

●David　まず、オープンサイエンスの主たる特徴は技術革新があったからこそ生まれたのだという考え方に、私は同意できません。オープンサイエンスはそれまでも存在していました。オープンサイエンスというのは、新しい物理学が登場した黄金時代に、物理学の国際化および世界会議の基礎となったものでした。古典物理学と量子力学の関係に関する論争は、オープンで行われていました。誰もがお互いのことを知っていて、会議で、あるいは論文上で、個人的に議論していたのです。それが物理の世界における知識の追求を行うに当たっての共同研究だと見なされていたのです。そしてそれは、国籍を超えて、誰にでも開かれた、科学の国際化でした。ただしそこには、ある一つの側面があって、それは共同での取り組みだったということです。少なくとも真理に近いものとして受け入れられるものを追求する共同研究でした。私たちは、決して真理に到達することはありません。まだ他に理解しなければならないことというのは、必ず出てきます。しかしながら、さしあたりは、今の段階で理解できていることで、そして今できる、その現象の観察結果を記録するやり方で、また、そのような現象を引き起こす仕組みを理解するやり方で、前に進んでいくしかありません。この場で議論されているのは、オープンサイエンスを、社会全体に優れた外部性をもたらすような形で強化していくことはできるのか、そしてそれを、平等主義的、民主主義的なやり方でできるのか、さらには、何を、どうやってしなければならないのかを決定する際に、スポンサーや研究者のパトロンたちでなく、いかに研究者たちを巻き込んでいけるかということです。

●フロア 2　ブーメランが返っている絵が描いてあるスライドを見て、私は非常に示唆的だと思いました。先ほどの質問と関係するのですが、一回 17 世紀に戻られています。それは、ルネサンスに戻る、ルネサンスというのは少し論理的に変かもしれませんが、はたはギリシャまで戻ろうかという話に結局なるわけです。ギリシャの人たちは自然科学ということはなく、あれは哲学であり、フィロソフィーです。それがアリストテレスにつながっていて、中世あたりのコペルニクスにしても、あれは完全にフィロソフィー、要するに論理の整合性で来るわけです。あれは論理の一貫性です。中世というのはそういう学問であり、それが 17 世紀になると実証的なものが加わってきました。オープンデータ、オープンサイエンスといわれている一つの影には、論理の代わりに、データから理解される現象の整合性というものをやっていこうというような観念があるのではないでしょうか。

もし 17 世紀にデータがたくさん、今のようにあったとすれば、多分、今のような自然科学があったと思います。自分でもよく分かりませんが、私が言いたかったことは、ディシプリンは哲学であるということです。今の世の中というのはディシプリンが完全に分化しているわけです。日本学術会議の何とか委員会などはその最たるもので、とにかくディシプリンになっています。日本の学問制度もディシプリンです。ディシプリン的な考え方が固定化されてしまっています。それを打破することがオープンデータなり、オープンサイエンスなりの流れで、それがディシプリンの境界を破壊できるだろう、そしてフィロソフィーを目指すのだということではないでしょうか。

ギリシャの昔に結局返って、何かの専門家ではなくて、いろいろな人がいろいろな現象、自然界のことをいろいろな分野から多角的に考えることができる、そういう文化をつくろうということなのだろうかと、少し妄言ですが、それを目指すのだったらこれは教育から変えなければ仕方がありません。今のように理・数・化・生・地など言っていては駄目で、そうなると、哲学者は音楽ということもあるのだから、音楽まで含めなければいけません。教育体制から全部変えないと、とても今言われているようなオープンサイエンス、いわんや、そういうものが人工頭脳でできるとは到底思えません。

やはりオープンサイエンスをやるのだったら、デー

タだけではなく、やはり教育に対する考え方を全部変えないと、物事は解決しないのではないかということを少し言いたかったのです。

●**村山**　今のお話と深貝さんの話は、ある意味新しい学術、サイエンスがあり得るかどうかというところに踏み込もうとしたのだと思うのですが、科学は 17 世紀から 20 世紀にかけてできたものがずっと現に続いていて、それを変える話ではないと思うのです。メディアが変わったことでオーサーシップは変わるかもしれません。ディセミネーションのやり方は変わるかもしれません。パブリケーションが変わるかもしれません。けれども、結論の導き方、パズルの解き方は変わらないと思っているのです。

例えば物理学ベースで全てが語れるかどうかは分かりませんが、一つのイグザンプルとして物理学を出すと、かつてはやはり理論ベースでした。数学でなければ物理学としての結論を導けないというのが、いわゆる 20 世紀の物理学帝国だったわけです。それが、データがあれば何とかなるぞという話が出てきたときに、本当にそれで解けるのかというと、実は解けません。論文にしようと思うと、データから導かれたものが過去の物理描像と整合するかどうかをきちんとチェックしなければ、ただ、データが現象を表現したからというのは、私の業界で言うところの、いわゆる報告しかしていない論文になるわけです。ですから、きちんと理解をしていないことになります。

ですから、科学的理解という意味では、今のデータの議論は、データの比重がものすごく増えたのにきちんと使えていない現状を何とかするための振り子の片一方であって、そこでデータが使い切れていないので、それをきちんと使いましょう。でも、そのときにもう片一方である、理論（theory）や数学（mathematics）を抜きにしていいのではなくて、データで十分知識を得尽くしたら、今度はもう一度そこから理論や数学に戻る可能性が十分にあると思います。その両方がないと理解になりません。

先ほど申し上げたような、人類にとっての将来を占うべき、曖昧だけれどもそれしかない、一番もっともらしい将来指針となる論理構築手法としての科学というものは、そういうデータという基礎資料に基づいて理屈で結論を導き出すので、両方がなければいけません。ですから、そこは、モードチェンジはないだろうということで、コミュニケーション部分がどれだけ物事を変えていくかというのが将来の、今後の課題だろうと思います。

●**深貝**　多分、今のお二人のうち、私はどちらに近いかと言われると中途半端です。というのは、オープンサイエンスはボーダーレスサイエンスではないのです。そういう意味で村山さんの言うとおりだと思います。

今までのように得られたデータを研究室で囲って分かったと言っているようなことではなくて、得られたデータはいったん明るみに出した方が、解釈の可能性も同じ学術コミュニティの中で広がるかもしれないし、思い込みをチェックできるかもしれない。学術のために研究予算を投じている中で似たり寄ったりの実験をあちらでもこちらでもやってうまくいかないのなら、手分けしてやってうまくいく可能性を高める方に予算が使われている方が国民にも説明しやすいなど、いろいろなことは確かにあると思います。そういう意味で、オープンなデータになって解釈にできるだけ多くの人が加わるということの方が望ましいわけです。従来型のディシプリンの中においても、多分、そういうことではないかなと思います。

●**村山**　そういう可能性を広げるということですね。

●**深貝**　はい。しかし、その場合に、問題なのは今の評価システムとどこかがかみ合わないことです。先に誰かが投稿して、それが早くレフェリーを通ったからという今までの文化とずれてしまうという問題はあると思います。

学術情報をマネージするという図書館レベルの話を

含めてどなたか、質問やご意見はいかがでしょうか。

●David　哲学に関するお話ですね。実は先ほどフロアから発言された先生（フロア 2）とは、私がずっと若かったころに出会いました。先生は学部のトップを何年も務めていらっしゃって、非常に高く評価されていました。ここで再びお会いできて、その変わらぬクリエイティビティに触れ、刺激的な発言をお聞きできて非常にうれしく思います。ディシプリンというものが時代遅れなのか、今日においては崩壊してしまっているのか、そしてその原因はデジタルテクノロジーにあるのかという問題ですね。

私は、それについては別の角度から、子どものころに自然に学ぶ概念ではなく、より難解な概念を扱う領域に対応していくための力を身に付けさせる際に、ディシプリンが果たす役割について考えたいと思います。そうした難解な概念の意味や意義を学ばなければならないとなると、そこには教育的目的が不可欠です。そして、ディシプリンの枠組みの中で系統立てて指導するということが、非常に生産性の高い営みであることは、既に分かっています。多くの人たちが、そういったやり方での教育を受けています。また、難解なテーマの教育を受けた人たちのほとんどは、たとえそれが、自分たちが何を受け入れたくないかを明確に定義するという形であったとしても、実際にその教育の恩恵を受けているのです。これは重要な点です。

方法論の変化によって、古典物理学のある側面は損なわれてしまいました。その側面とは、数学的な定式化や解法に頼ることで、システムが平衡に向かう傾向を見いだし、私たちが何を永続的に観察するのかを説明するというところです。例えばそれがカオスの世界であれば、その現象を安定していると言い表すことはできません。

数学的解法という方法論が高く評価されるようになり、理論を定式化すれば手近な数学で解くことができるということで、それが主流となったのは、演算が不可能だったからです。システム全体を記述し、形式的に解くことは可能です。そうやって概念上ではできるのですが、実際に演算で解を導き出すことはできませんでした。理論家にはこの強みがあったので、理論が実験や観察より分野で優勢になったのです。理論家たちは、誰よりも優位な立場にありました。そして最終的には、天文学などにおける実験成果によって二つ目の柱が生まれ、二つの体系が出来上がりました。実験物理学と、数理物理学です。この二つの柱のうち、一方はもう一方よりも上でした。そして、上の方に座っていたのが、理論家たちでした。

コンピュータの革命によって演算は安価になり、演算が不可能だったものからもアルゴリズム的解法を得られるようになりました。これにより小さな革命が起こり、そこにいくつか良い出会いがあって、理論の枠組みが古典物理学から経済学へと引き継がれていきました。そして、それらのモデルの多くが、時間的変化などを表すモデルの主流となったのです。

数年前、サンタフェ研究所で、物理学者と社会科学者の会合が開催されました。まず、物理学者であるアンダーソン教授が、物理学者は何をやっているかという話をし、その後に故人となられたケネス・アロー教授が、経済学者は何をやっているかについてプレゼンテーションを行いました。このアロー教授の話が終わるころに、アンダーソン教授がこう言ったのです。「なぜあなた方は、そうした数学の問題を解いているのですか？　私たち物理学者はもう定理を証明したりはしていません。何もかもを演算しています。現象のシミュレーションを行い、それを演算したり調整したりしています。そうやって前に進んでいるのです」と。そして残念なことに、アンダーソン教授はこう続けました。「それは良いことでした。というのも、われわれは今や量子電磁力学をやっているのですが、この量子電磁力学を展開した第一人者は、より効率的に演算を編成できるということを示してくれたからです。最初はコンピュータを使わずに」。そしてそれが大いなる進歩への扉を開き、そうした方法を用いれば、洗練されたアルゴリズムの演算を行うことができるように

なりました。そんなわけで、この後退したり前進したりしている変化があって、そこではテクノロジーが重要な役割を果たしています。

そしてこれはガリレオが抱えていた問題でもありました。ガリレオは月を観察し、そこに斑点があることに気がつきました。しかし懐疑派たちはこう訊ねます。「どうしてその斑点が望遠鏡のせいで見えているものではないと分かるのですか？　結局のところ、あなたの望遠鏡を使わなければ、あなたの発見を検証することはできません。私が別の望遠鏡を作ったとして、同じ答えが得られるでしょうか？」。適切な望遠鏡さえ作ることができれば、ガリレオの発見を検証できるというのが、その答えです。同じ人物に望遠鏡を作ってもらう必要はありませんが、例えば光学などに関しては同じレシピに従う必要があります。そうして彼らも同じものが見ることができて、論争には決着がつきました。

●**武田**　例えば 16 世紀から 17 世紀にかけて数学がサイエンスの価値をつくり、サイエンスのお客さんをつくり、有用性をつくったことによって科学が発展したということはよく分かります。逆にそれが今何なのかと言うと、今はまさにコンピュータがあり、先ほど村山さんは物理的に説明がつかないものはクズだと言っていましたが、実は今はそうではなくて、ディープラーニングで解決したらそれでいいですという人たちがいるわけです。実際に有用性が証明されているわけです。それはもしかすると別の、われわれが言っているサイエンスではない別のサイエンスをつくる可能性があります。数学が今のサイエンスをつくったならば、ディープラーニングのようなコンピュータ、つまりアルゴリズムに依存するものが別のサイエンスをつくる可能性があります。そのあるかないかの議論がしたいわけではなく、それをサイエンスと呼ぶかどうかは別として、そういう世界が今、われわれの今の科学のコミュニティの外にできつつあるのです。

ここが私が一番気になるところです。頭の良いソフトウェアデベロッパーが、非常に良いアルゴリズムを書くわけです。もちろんサイエンスのディシプリンで勉強したかもしれませんが、彼らが活躍する場は科学の場ではなくてビジネスの場で、それでお金をもらいます。そういうエコシステムが今できつつあり、彼らは論文を書く必要はありません。だから、今、GitHub にはたくさんアルゴリズムペーパーが置いてあります。彼らは科学のレピュテーションは要らないのです。アルゴリズムを置いておくから、それで GitHub にソースも置いてあるのでバリデーションは可能なわけです。でも科学のレピュテーションは要らない、そういう世界が今できつつあります。

この世界を皆さんはどうしたいと思いますか。質問としては、これを科学の世界に取り込みたいですか、それとも「いいです、勝手にやってください」としたいですか。この質問を答えられる方にぜひ答えていただきたいのです。

●**倉田**　私は取り込みたいというか、それを取り込むものとしてイメージしているのが、私の発表で言った Scholarly Commons です。つまり、民間の研究者も含めた形で、何かの現象に関してある特定の目的を持って、同じ枠組みやメソドロジーを持ってやれる人たちが集まれるコモンズがあったらいいなと思います。私は、将来的なスカラリーコミュニケーションのエコシステムとはそういうものであってほしいと願っています。

●**David**　確かに今、アルゴリズムをトレーニングして、意思決定をさせたり、顔のマッチングをさせたりというマシンラーニングがブームになっています。アルゴリズムを使えば、多くのピクセルを取り込み、異なる二つのものを区別することができるのです。この考え方は部分的には、人間はパターンを認識することで学習できるというところから来ています。データや観測結果を、特定の反復的空間パターンが観測できるような形で表示してやれば、そこから学習できるとい

うことです。元々の問いは、人間は、表示されたパターンから、興味をひく安定した付帯現象を導き出すことができるのかというものでした。すなわち、非常に多くのデータを取ってきて、それを操作したり、変換したりして、何か目立った発見をすることができるのかというもので、ここでの問題は人間の学習です。人間には知覚の限界があるが、それと同時に、非常に初期の段階で学ぶパターン認識能力も備わっているということもあり得るのでしょうか？　私たちが考えもせずに下している決断の多くには、パターン認識が関わっています。例えば、まばたきの反応や、何か物体が飛んでくることで首をすくめてしまうといった行動、これは処理をせずに行われています。古くなった脳には、多くの事柄がエンコードされているのです。

だからこそ私たちは、特定の歴史的従属物を認識することができます。つまり問題は、私たち人間は、自分たちに物事を見せてくれるマシンを使いこなしていけるのか、そして自分たちが興味を持つであろう物事を認識できるかどうかというところです。というのも、マシンは明確な特徴やパターンを持っています。これはマシンを扱う人間の限界です。一方で、マシンラーニングというものがあります。マシンは人間と一緒にやっていけるのでしょうか？　マシンラーニングは、シミュレーションとは全く違うものなのでしょうか？基本的に、マシンラーニングの場合は、マシンに人間の顔をどうやって認識して何をするのかを学ばせて、こちらが思うように行動してくれるように訓練してやる必要があります。つまり、自分自身をマシン上に再現するのです。そうすると、自分自身の限界がマシンに伝わることになり、ある意味で、自分と同じ程度の賢さや馬鹿さを備えたマシンが作られることになります。それは、マシンが「適切に」行動するように人間の知性を使わないと、訓練することはできないからです。AI はそうした問題を解決してくれません。この問題はずっと残っていくでしょう。

●フロア 3　内閣府総合科学技術・イノベーション会議議員です。Paul David の考え方を私なりに説明させていただきたいのですが、人間は五感を使って自然現象をオブザーブしてきました。それに対して、どんどんいろいろな道具を獲得してきました。多分、AI もその道具の一つであるという認識です。その中で、五感で見えること以外に、見逃していることもたくさんあるのです。データを AI に通すと、人間が見逃すような特殊性、人間ではキャプチャーできないものが現象として浮き出てくるということがあります。

逆に言えば、これまでは頭の中で仮説をつくってから実験をしたのですが、そこから仮説を引き出すことも今は可能になっています。そういう意味で、AI サイエンティストというものも出てきています。それはある種のルーチン化された実験をやらせてしまえばいいというところまでいっています。その先で、人間がどういう知恵を使ってその次のところへ行くか。最初の話に戻るのですが、まさにそこから、これは面白い、不思議だと思うことを、人間が感性でキャプチャーして、それを今度は実験のプロトコルに落としていって仮説をつくっていく。そういう意味で、行ったり来たりのインタラクションがあるから面白いのです。

でも、気を付けなくてはいけないのが、全部を AI に任せきりにしてしまうと、人間は何なのか、サイエンティストは何なのかという話になることです。ですから、すごく面白いことが今起こっているので、それをうまく使い込んでいくトレーニングを教育の中でもしなければいけません。

今、私が一番危惧しているのは、私たちの世代は、大人になってからコンピュータなどいろいろな道具を使うようになりましたが、今の子どもたちは初めからその環境の中で育っていくことです。何かふっと感じるところとか、「風が吹いていい気持ちだ」とか、「これは何？」というキュリオシティというものを、バーチャルではないリアリティの中で感じることを体験する機会を意図的につくっておかないと、やはり不完全な人間になってしまいます。その辺を問題意識として持っています。

●**深貝**　ありがとうございます。まとめとしてモデレーターとして言わなければいけないことの半ばほどは今お話ししていただきました。

20 世紀半ばには、生物の、例えばアリなどの社会がどうなっているかを見て、人間のモラルのヒントがないかと探っていましたが、今度は人工知能がそのモラルのヒントをくれるかもしれません。それは科学的ではないかもしれないけれども、振る舞い方は役に立つかもしれないという意味で有用性はありそうです。ただ今、先生が最後におっしゃったことで言えば、情報端末の教科書を使った学校教育が始まっていますが、5 年ぐらい先にはそういう学生が大学に入ります。背表紙がある本が逆に珍しいという世代に、図書館は何をするかということまで含めて待ち受けています。しかしそれは、20 世紀初めまでは歩いたり馬に乗ったりしていて自転車が珍しかったのが、今はもう車社会になってしまったのと同じくらいのことだと思えば、どうということはないかもしれません。いずれにしても今日は、これからの知識がどのようになっていくかということを考えていく大変いい機会になったかと思います。

パネリストの皆さま、いささか禅問答のような質問ばかりしまして、大変失礼しました。どうもありがとうございました。

●**林**　モデレーターの深貝先生もどうもありがとうございました。これでパネルディスカッションのセッションを終わりにしたいと思います。

国際学術情報流通基盤整備事業（SPARC Japan）年報
―平成 29 (2017) 年度―

平成 31 年 3 月

発行　大学共同利用機関法人　情報・システム研究機構
国立情報学研究所　学術基盤推進部　学術コンテンツ課

〒101-8430　東京都千代田区一ツ橋 2 丁目 1 番 2 号
TEL　03-4212-2351
FAX　03-4212-2375
E-mail　sparc@nii.ac.jp

CC BY