

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。

※所属、肩書はすべて開催当時のものです。

CONTENTS

■ SPARC Japan 活動報告

学術情報流通推進委員会

学術情報流通推進委員会について

■ 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・パネルディスカッション)

■ SPARC Japan 活動報告

学術情報流通推進委員会

学術情報流通推進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/>



学術情報流通推進委員会について

我が国の学術雑誌の電子化・国際化への対応等を課題として、2003 年に国際学術情報流通基盤整備事業 (SPARC Japan) が始まりました。2003 年以降、学術情報流通における潮流の変容に応じて、本事業も内容の見直しを行ってきましたが、第 5 期の終了 (2018 年度) にあたり、特定の事業ではなく、ステークホルダー間の連絡調整を行うことによって、学術情報流通基盤整備を推進するという役割に大きく舵を切ることとなりました。

そこで 2019 年度から、活動方針を策定する主体の名称を、国際学術情報流通基盤整備事業運営委員会から学術情報流通推進委員会へと変更し、次の基本方針に基づいて活動を行うこととなります。

近年の情報通信技術の進展に伴い、学术论文のオープンアクセスに加えて、研究データを含めた研究プロセスのデジタル化と共有に取り組む、オープンサイエンスが国内外で進展しつつある。学术论文や研究プロセスの相互利用の促進は新たな知の創出にも資することから、学術情報流通推進委員会の第1期においては、オープンアクセス、オープンサイエンスを推進するために、国内外の学術情報流通の動向や実態の把握に努め、それらに基づいた学術情報の公開や利活用に係る戦略の検討と調整、アドボカシー活動等を、学術コミュニティ等を中心としたステークホルダーの参画や連携のもとに行う。

学術情報流通推進委員会はこの基本方針に則って、第1期(2019年度～2021年度)においては、(1)国内ステークホルダーとの協調、(2)国際協調に係る戦略の検討と提言、(3)アドボカシー活動の実施、(4)学術情報流通の動向に係る調査の提言に重点的に取り組みます。活動の詳細については、適宜 web サイト等に掲載する予定ですので、ご参照ください。

■ SPARC Japan セミナー報告



第2回 SPARC Japan セミナー2018

「オープンサイエンス時代のクオリティコントロールを見通す」

2018年10月25日(木) 国立情報学研究所 12F 会議室 参加者：53名

今回のセミナーでは、オープンサイエンス時代のクオリティコントロールの方向性とコンテンツの質の保証をテーマに、現状の具体的な試みの情報共有と議論を行いました。

2017年度に開催したセミナーにおいては、プレプリントサーバや機関リポジトリに収録されるコンテンツの質の担保が論点となり、また、サイエンスにおける普遍的な課題としての質の担保の必要性が再確認されました。今回は、F1000Research におけるピアレビューやバージョン管理の実践、arXiv や BioRxiv といった伝統的なジャーナルへの投稿・掲載といった方法以外のプラットフォームにおける研究成果公表の実践を取り上げた他、研究者の視点から、公表手段を選択する際の条件や様々なプラットフォームの活用に関する現状、そしてそれらに関連する課題を紹介いただきました。

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/2018/20181025.html>



企画概要

オープンサイエンスは科学のあり方そのものを変容させることを予見しており、学術雑誌の電子化にとどまらず ICT を活用した様々な成果公開の試みなどがすでに取り組まれている。その際、オープンに公開される情報の質が保証されることで、安心して利活用が進み、オープンサイエンスも進展する。学術雑誌が電子化される前から質の保証の手段である査読(ピアレビュー)のあり方は常に議論されていたが、電子化とオープン化によって、オープンピアレビュー、軽量査読など質の保証についても様々な試みが生まれている。最近では、コンピューターサイエンス等の分野によっては、プレプリントに上げること自体が主な研究成果公開手段となりつつある中で、どのようにその成果を評価するかも喫緊の課題となっている。あるいは、英国の研究助成団体である Wellcome Trust が、査読付きのオープンな研究論文、データ出版プラットフォーム (Wellcome Open Research) を運営するようになっている。



このような状況下、2017年度の第2回 SPARC Japan セミナーではプレプリントサーバ・機関リポジトリの取組を議論する中で、質の担保が論点となり、議論を掘り下げる必要性を強く認識した。さらに、第3回ではサイエンスの成り立ちとピアレビューの始まりが語られ、コンテンツの質の担保は古今を問わない普遍的な課題であることを再確認した。ICTの発展とオープン化が新たな研究環境を生み出している中、特定の研究者コミュニティ内のピアレビューを超えて、分野間、あるいは、市民を含む誰もが科学研究に参加できる場において、質の保証を行う仕組みをどう開発していくかは今後の学術コミュニケーションを見通す上で非常に重要なポイントとなる。

そこで、本年度第2回の SPARC Japan セミナーでは、オープンサイエンス時代のクオリティコントロールの方向性とコンテンツの質の保証をテーマに、現状の具体的な試みについて最新の情報共有と議論を行う。

参加者から

(大学／大学・教育関係)

- ・学術出版の新しい方向性を知ることが出来た。

(企業／学術誌編集関係)

- ・研究者の方がプレプリントについてどう考えているのか分かり、大変参考になりました。

(企業／その他)

- ・内容はとても刺激的でかつ新しく、また自分にとって研究者のリアルな行動を知られて大変役に立ちました。

(学協会／学術誌編集関係)

- ・オープンサイエンスの潮流、研究者の立場、質の担保についてよく理解ができました。こういった学術分野の最新情報を Web 中継を通じて共有していただけたことに感謝いたします。〔動画中継視聴者〕



企画後記

😊 企画主旨の裏側の話として、2017年1月に Wellcome Trust を訪問して創設間もない Wellcome Open Research について担当者と議論し、翌年4月の OECD のワークショップにて Rebecca さんと一緒に登壇したことで、この企画の実質的な下ごしらえができました。難しい議論でしたが、世界的にも明快な答えがまだないテーマについて、日本においても主体的に考えたことを記録できることが重要と考えています。

林 和弘

(科学技術・学術政策研究所)

😊 本年度から企画 WG に参加しました。研究者、出版者としてオープン化をどのように推し進めていくべきかとい哲学に裏打ちされた議論がなされ、運営に携わる側も刺激を受ける内容でした。楽屋話ですが、Rebecca さんに忙しい会議スケジュールの合間を縫って行ける観光スポットを尋ねられ、数名で悩んだということがありました。江戸・東京研究にも携わる身として、研究から日常にまで使える情報もオープンにせねばと改めて実感しました。

鈴木 親彦

(国立情報学研究所 /

人文学オープンデータ共同利用センター)

😊 初めて企画立ち上げから携わせていただきました。従来の商業出版による学術情報流通にノーを突きつける Rebecca さんや、実際にプレプリントサーバを利用する研究者の方々のお話を間近でうかがうことができ大変刺激的でした。研究の質の担保に図

書館員が関わってよいのかという点が企画段階から難問だったのですが、パネルディスカッションでデータ管理という指針が示され、その重要性を再認識しました。

中原 由美子

(筑波大学 学術情報部)

😊 今年度から WG メンバーとなり、はじめて企画立案から担当した回でした。プレプリントの最新動向、オープン化時代の研究の質保証といった、図書館の実務とは少し (いや、とても?) 遠いテーマだったので、一番勉強させてもらったのは私かもしれません。これを契機に図書館がどう関わりうるのかを考えてみたいと思います。

中村 美里

(東京大学附属図書館)

😊 いろんな意味でチャレンジの多い企画でした。まず、テーマ。今回はクオリティの担保ということで、従来とは90度異なる視点でオープンサイエンスを捉えることができました。次に、手法。遠隔地にいる登壇者が Web 会議システムを使って講演を行いました。事務局は準備が大変だったと思いますが、海外の方を含めてより多くの方にご登壇いただけるチャンスが広がりました。今後も SPARC Japan セミナーはたくさんのチャレンジに挑んでいくことを期待します。

八塚 茂

(バイオサイエンスデータベースセンター)

本誌についてのお問い合わせ

国立情報学研究所 SPARC 担当

E-mail co_sparc_all@nii.ac.jp FAX 03-4212-2375

<https://www.nii.ac.jp/sparc/>

第 2 回 SPARC Japan セミナー2018

「オープンサイエンス時代のクオリティコントロールを見通す」

開会挨拶 / 概要説明

八塚 茂

(科学技術振興機構 バイオサイエンスデータベースセンター)



八塚 茂

国立研究開発法人科学技術振興機構バイオサイエンスデータベースセンター研究員。システムエンジニア等を経て、2015年10月より現職。研究対象が多様でかつ各研究機関に分散している生命科学系のデータを収集し、調査や整理を行った上でメタデータを付与し、明確な利用許諾のもとで公開する事業に携わっている。



本日は、第2回 SPARC Japan セミナー2018「オープンサイエンス時代のクオリティコントロールを見通す」にご参加いただきまして、誠にありがとうございます。今週10月22日から28日までは世界的なオープンアクセスウィークです。本セミナーは、この一環としての「オープンアクセス・サミット 2018」というイベントにもなっています。

オープン化のアクセラとブレーキ

今年度の SPARC Japan の年間テーマは、「オープンサイエンスの定着に向けて」です。オープンアクセス、オープンサイエンスなど、似たような言葉が多く飛び交いますが、要は、サイエンスはオープンであるし、オープンにしていくことが大切なのではないかと理解しています。

前回の第1回セミナーは、「データ利活用ポリシーと研究者・ライブラリアンの役割」と銘打って行いました。こちらでは、データオープン化を進めるための利活用ポリシーやオープン化を担う人材としての研究者・ライブラリアンに焦点を当てました。これはオー

ペン化のための、いわばアクセラの部分であると理解しています。

一方、今回のセミナーでは、オープン化されるもののクオリティコントロールに焦点を当てています。これはいわばブレーキに当たるものだと考えています。もちろん、ブレーキといっても、オープン化を阻害するものではなく、むしろオープン化を前に進めるための必要な機構であると考えています。

本日は、オープンサイエンス時代のクオリティコントロールはどのようなものであるべきかということにつきまして、皆さまと共に考えていきたいと思います。

セミナーの流れ

本日は、午前中に2件、講演を行います。これはいずれも英語です。その後でいったん Q&A セッションを行い、会場の皆さまからのご質問を受けたいと思います。午後も2件の講演があり、こちらは全て日本語です。その後でパネルディスカッションに入ります。大変な長丁場になりますが、皆さまの積極的な議論へのご参加をお願いいたします。

第2回 SPARC Japan セミナー2018

「オープンサイエンス時代のクオリティコントロールを見通す」

ジャーナルを超えた動き： 出版者、資金提供者、機関の変わりゆく役割

Rebecca Lawrence

(F1000)

講演要旨



研究ジャーナルにおける新しい知見を出版する昔ながらの手法は時代遅れで高コスト、持続不可能になっており、科学の進歩を遅らせたり、ダメージを与えたりしている。更なるオープンアクセスやオープンサイエンス政策への世界的な著しい転換は、技術革新も相まって、今や多くの課題を処理することを可能にしつつある。2013年に、F1000は世界初のオープンリサーチ出版プラットフォームであるF1000Researchをスタートし、迅速な出版の能力と、研究におけるより高い透明性、堅牢性、そして再現性を確保するという機能を結びつけた。私たちのアプローチでは、ピアレビューによる透明性と論文のバージョン管理ということによる、FAIR原則に基づくデータシェアリングとともにほぼ即時の出版を行っている。これにより、管理ということを著者らのもとに戻し、伝統的でひどく時代遅れになってしまった研究出版の様式によって引き起こされる多くの課題を解決することを目指している。これは、研究を行い、影響力をもたらすための、より高い透明性があり、協力的で適切な方法への転換を容易にすることになる。

オープンリサーチ出版への沸き起こる関心に対して、また実証可能な利益ということについては、いま我々が世界中の、数多くの人目をひく研究資金提供機関や研究機関、例えばビル&メリンダ・ゲイツ財団、ウェルカム・トラスト、アフリカ科学アカデミーに出版プラットフォームを提供していることで証明できる。このアプローチはエコシステムにおける出版者、資金提供者、機関のそれぞれの役割を変え、最近の研究や研究者の評価システムに関するよくある課題の多くに最終的に対処する機会になる。

Rebecca Lawrence

レベッカ・ローレンスはF1000のマネージングディレクターであり、新しい科学的知見のライティング、出版、発見、評価の面で、研究コミュニティをサポートするツールやサービスの提供を行っている。また、新しいオープンサイエンス出版プラットフォームであるF1000Researchの立ち上げの責任者でもあった。ほかにも、Wellcome Open Research、Gates Open Researchなど多くの資金提供者や機関ベースの出版プラットフォームの立ち上げの影でイニシアチブをとってきた。この取り組みは、科学的知見とデータが伝達され、結果として研究や研究者が評価されるという場において新しい軌道を生み出すという目的がある。さらに、欧州委員会のオープンサイエンス政策プラットフォームのハイレベルアドバイザリーグループのメンバーであり、次世代の指針作りや、助言の取りまとめを行うOSPP-RECの議長である。また、データやピアレビューに重点を置く、例えばRDA、CASRAI、ORCIDといった機関のいくつかのワーキンググループの共同議長も担ってきた。データポリシーや規則のイニシアチブであるFAIRsharingと、DORA (San Francisco Declaration on Research Assessment) のアドバイザリーボードメンバーでもある。彼女が築き、育てたDrug Discovery Groupのあるエルゼビアを含むいくつかの出版社で、約20年に渡りSTM出版に関わってきた。もともとは、薬剤師としての教育を受けて資格を取っており、心血管薬理学の博士号を取得している。



既存の出版システムの課題

既存の出版システムには多数のさまざまな課題があります。一つ目は、新たに研究で得られた発見の大部分がオープンアクセスでないということです。依然として購読料の壁に阻まれています。

もう一つの問題は、研究者が何かを発見した時点と、その発見を他の研究者や世の中と共有する準備が整う時点、そして他の人たちが実際にその研究を目にする時点の間に大きなタイムラグがあることです。その時間差は数カ月から数年になる場合も多く、実際のところ、正当な理由があってもこのような遅れが生じているわけではなく、この遅れは著者、読者、利用者にとって、何のメリット也没有ありません。

ほとんどのジャーナルは匿名の査読を行います、これはつまり、何が行われたのか、誰が論文を査読したのか、どのような決定が下されたのか、なぜエディターがその決定を下したのか、分からないということです。従って、これは対立やバイアスが内在する仕組みになっており、それは必要のないものだとも私たちは考えています。

次に、研究のほとんどは、基礎データが公表されることなく出版されています。データがなかったら、研究で得た発見をどうレビューし評価すればいいのか、理解するのがとても困難です。データがなかったら、新しい発見を試したり、再現したり、再利用したりすることもとても困難です。

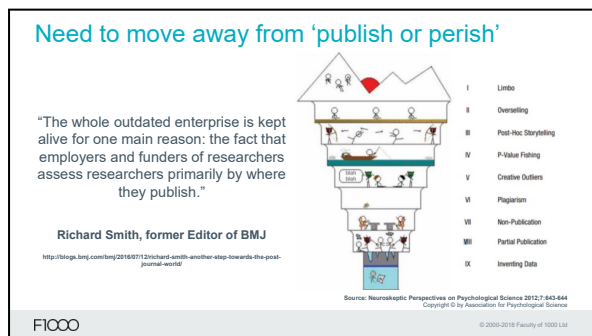
それから、全く出版されない優れた研究もたくさんあります。ネガティブな発見や追加的な発見などです。ジャーナルはこのような研究を掲載しがらないため、

出版されません。投稿数を左右するジャーナルのインパクトファクターは引用数に左右されますが、このような研究はあまり引用につながらないからです。そのため、ジャーナルは興味を示しません。試算によると、優れた研究のうちほぼ半数が一度も出版されていないのです。

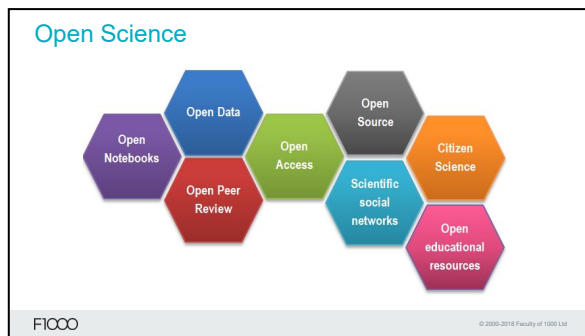
そうすると、ポジティブな発見だけ明らかにされ、ネガティブな発見は知られないので、科学に対する私たちの理解が歪んでしまうだけでなく、研究が繰り返行われたり、不必要な資金提供を受けることになったりして、深刻な研究の無駄が生じます。

ですから、私たちは「パブリッシュ・オア・ペリッシュ（出版せよ、さもなければ死せよ）」のシステムから移行する必要があります。このシステムは、研究者による自らの発見の過大評価や後付解析、p 値ハッキング、さらには一部のケースにおける盗用、データ捏造による不正など、さまざまな問題を引き起こしています（図 1）。『ブリティッシュ・メディカル・ジャーナル』の元編集者であるリチャード・スミスは、このように言っています。「時代遅れのこの事業全体が、一つの理由のために生かされ続けている。研究者の雇用主と資金提供者が、研究者を主に論文の掲載誌で評価するという理由だ」。

オープンサイエンスは、このような問題の多くに対処することを目指しています。オープンサイエンスには多数の側面があって、オープンアクセスというだけではありません。オープンサイエンスはよくオープンアクセスと混同されがちですが、オープンアクセスに加え、オープンなデータ、オープンな査読、オープン



(図 1)



(図 2)

なノートブック、市民科学といった側面を持ち合わせています。あらゆる問題がオープンサイエンスを成しているのです（図2）。

世界的なオープンサイエンスへの移行

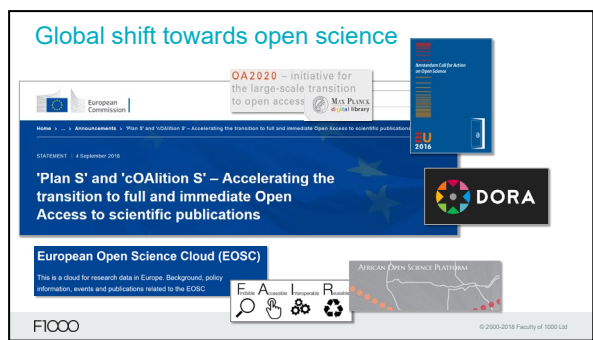
既にご存じのとおり、オープンサイエンスに向けた世界的な移行が起こっています。世界中で政策決定における変化が起こっていますが、特にヨーロッパで顕著です。欧州委員会は、2016年にオープンサイエンスに関するアムステルダム行動要請（Amsterdam Call for Action on Open Science）を発表しました（図3）。これにより、オープンサイエンスシステムに移行するための政策を実施し、変化を促す、多数のイニシアチブが生まれました。オープンアクセスに向かうためのイニシアチブが多数あります。Max Planck Digital Library は、OA2020 というイニシアチブを立ち上げました。これは、購読型のジャーナルをオープンアクセス型のジャーナルに転換させ、各機関が購読する際に、その購読パッケージにオープンアクセス料も組み込むよう奨励することを目指すというものです。

Plan S は、近年で最も重要なイニシアチブの一つです。これは欧州委員会と欧州研究会議が他に 11 の資金提供者と共に進めているもので（cOAlition S は成長を続けています）、これらの資金提供者は 2020 年以降速やかにオープンアクセスにすることを求めると述べています。まだやらなければならないことはたくさんありますが、これはオープンアクセスに向けた重要な変化です。DORA も重要なイニシアチブです。これはサンフランシスコ研究評価宣言（San Francisco Declaration of Research Assessment）のことで、2012年に立ち上げられました。多数の機関と個人がこの宣言に署名し、研究と研究者の評価方法をインパクトファクター以外のものに変更する方針であることを訴えました。しかし、そのためにやるべきことはまだまだたくさん残っています。

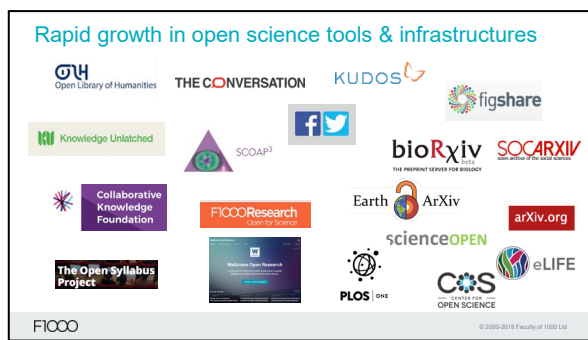
データに関しては、世界中、米国や日本、他の場所にも、いろいろな重要なイニシアチブがあります。主なイニシアチブには、データを「見つけられる（Findable）」「アクセスできる（Accessible）」「相互運用できる（Interoperable）」「再利用できる（Reusable）」ようにする FAIR イニシアチブや、欧州オープンサイエンスクラウド（EOSC）、アフリカ・オープンサイエンス・プラットフォーム（AOSP）などの、データ共有やデータサービスを可能にするプラットフォームがあります。

また、オープンサイエンスのツールとインフラも急速に成長しています（図4）。例えば、近年多数のプレプリントサービスが立ち上げられています。そしてもちろん、物理学の分野で古くから存在する arXiv は、サーバー上で迅速に新しい発見や論文を共有できるようにしています。技術的な査読のみのプロセスに移行した最初のジャーナルである『PLOS ONE』や、協力型査読アプローチを取る『eLife』などの新しい出版モデルも多数あります。

また、私たちが研究について議論し、コメントする方法においても変化が強まっています。ただ単に新しい研究について知らせるだけでなく、それについて議論したりディベートしたりするために、Twitter や



(図3)



(図4)

Facebook、ブログ、あるいは The Conversation などのツールを使用するケースが増加しています。Collaborative Knowledge Foundation (Coko) などの新しいインフラプロジェクトも多数あります。また、SCOAP³、Knowledge Unlatched、Open Library of Humanities といった、さまざまな機関がまとまってオープンアクセスの費用を負担し、著者が出版に必要な資金の獲得の心配をしなくていいようにするアプローチも多数あります。これらは全て素晴らしいもので、新しいアプローチにおける変化のスピードが加速していますが、研究者がこれらの新しいアプローチを利用できるよう、完全にオープンサイエンスに移行することが依然として課題になっています。

オープンサイエンスの普及に対する主な障壁

オープンサイエンスの普及にとって一番の障壁になっているものは何でしょうか。最大の障壁の一つは、研究者がいまだに、インパクトファクターや論文を出版するジャーナルのブランドで評価されているということです。インパクトファクターやジャーナルのブランドは、評価システム全体に深く根付いており、とても使い勝手がいいため、どうすればこれを変えていけるかという、とても難しいです。評価というのは複雑な問題ですから、同じぐらい使いやすい代替措置というのは、やはり同じくらい欠点のあるものになってしまうので、何を評価しようとするのかということに基づいて適切なメトリクスを使用しなければなりません。また、オープンサイエンスは良質な科学ではないという誤解もあり、良質な科学もオープンサイエンスになり得る、オープンサイエンスも良質な科学になり得ると理解することが重要です。オープンで、良質な科学でなければならないのです。

また、現場の実情という課題もあります。イギリスの Wellcome Trust のような一部の資金提供者は、出版の場ではなく研究そのものに基づいて研究を判断すると以前から言ってきました。重要なのは、審査グループの現場において、このようなポリシーの遵守を保証

することなのです。審査グループの個人に話をすると、「それはそうですが、『Nature』や『Cell』、『Science』などで出版しているのが目に入るので、評価にそちらを使う方がずっと楽なんです」とよく言われます。また、大学のランキングの評価基準に至るまで、あらゆるレベルで変化が必要です。大学が所属研究者の出版物のインパクトファクターに基づいて評価されるのであれば、大学は所属研究者にインパクトファクターの高いジャーナルで出版するよう要請するでしょう。そうすれば、大学以外では評価に変化が生じていても、システム全体は変化しなくなってしまいます。

鍵：出版と評価を切り離す

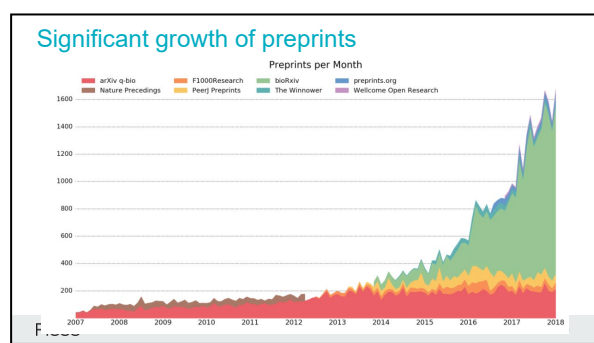
私たちの考えでは、このシステム全体を変える鍵は、出版に関する決定を内容の評価から切り離し、ジャーナルから離れることです。現在ではオンラインで出版していますから、ジャーナルは本当は必要ないのです。特に読者はジャーナルを必要としていません。読者は大抵 Google Scholar や PubMed を検索したり、その他のツールやアプローチを使ったりして論文を探しています。現在では、ジャーナルを必要とするのはジャーナルがキャリアに与えてくれるメリットを求める著者だけです。研究者が新しい発見をしたら、ゲートキーパーやエディターに却下されたり、私たちににとってはたいして面白くないと言われたりすることなく、コミュニティと共有できるべきです。同じく、研究コミュニティは、新しい発見を時間を置かずに閲覧できるべきです。読者や利用者は、新しい発見に対する同じ分野の専門家の見解に触れることで、真の利益を受けることができます。さらに、査読者はこのような発見を審査し、専門家としての意見を提供するという重要な仕事に対して当然の対価を得るべきです。新しい発見はその発見の質に基づいてのみ評価されるべきで、発表の場は関係ありません。

このような動きに向けた最初の重要な一步はプレプリントで、プレプリントサーバーで研究をすぐに共有できるようにすることです。これは特に生物学の分野

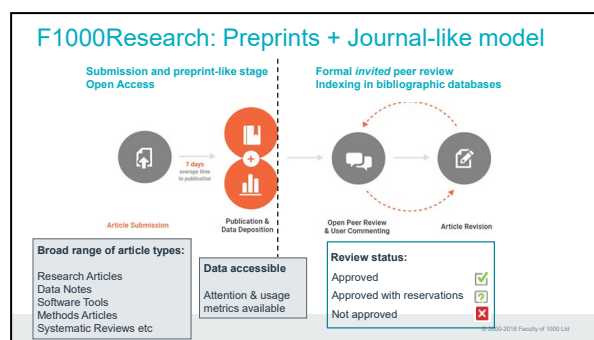
で非常に目覚ましい成長が見られています。例えば、bioRxiv では、ご覧のとおり投稿されるプレプリントが急増しています（図 5）。プレプリントは、有意でない結果、ネガティブな結果、追加的な結果を含め、あらゆる種類の発見を素早く共有でき、読者がこれらの発見にすぐにアクセスできるというメリットがあります。ただ、興味深いことに、ずっと前から同じことをしてきた物理学の研究者であっても、依然としてジャーナルで研究成果を出版しようとし続けています。その理由は、そうすればそのジャーナルのインパクトファクターを獲得することができ、キャリアの役に立つからです。

F1000Research

これは意味のないことです。ですから、私たちは F1000Research という出版モデルを開発しました（図 6）。これはプレプリントのメリットと、独立した査読、アーカイブ化、インデックス化、XML といったジャーナルのメリットを統合したプラットフォームです。まずは、プレプリントのようにします。著者はそのア



（図 5）

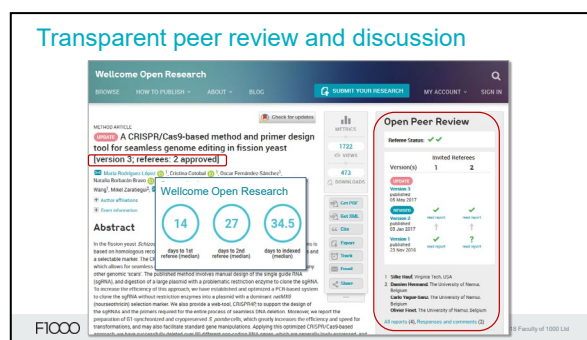


（図 6）

ウトプットに最も適したフォーマットであらゆる種類の論文を投稿することができます。投稿されたら、私たちの内部チームが一連の客観的なチェックを行います。私たちはこれが一つの研究であり、著者が一般に認められた組織に所属しており、盗作しておらず、結果を裏付けるデータがあるかどうか確認し、これらのチェックの結果問題がなければ論文を掲載します。これはとてもプレプリントに似ていますが、決定的な違いは、一度掲載されたら、ジャーナルに持って行って出版することはできないということです。ここで掲載されることがジャーナルに掲載されるということになり、掲載の時点で査読が開始されます。

また、著者は投稿する際に、私たちが著者に提供するリストの中から希望する査読者を複数名選択します。著者が他の査読者を提案することもできますが、私たちのチームがその査読者が関連分野の専門家であり、利益相反がないことを確認します。このプロセスの鍵となるのは、完全にオープンで透明だということです。論文と合わせて、査読者の名前と査読コメントも掲載されます。もう一つの鍵は、エディターがいないということです。これは著者が主導するプロセスで、プロセス全体を通じてどうしたいか決めるのは著者です。

掲載時に、私たちは著者に代わって専門家である査読者に査読を依頼し、査読者は二つのことをします。査読者は、査読レポートをまとめ、査読ステータスを提示します。三つの選択肢があり、「承認」（緑のチェックマーク）、major revision にあたる「条件付き承認」（緑のクエスチョンマーク）、「非承認」（赤のバツ）です。「非承認」というのはもちろん、既に掲載され



（図 7）

ているわけですから、リジェクトという意味ではありません。ただ承認されていないということです。私たちは、PubMed や Medline、Scopus などの重要な文献インデックスと合意し、論文が「承認」を二つ、または「承認」一つと「条件付き承認」二つを獲得したら、インデックス化すること、また、今後と過去の全てのバージョンもインデックス化することを決めました。

実際の論文で見てみる方がこのプロセスを理解しやすいでしょう（図 7）。ご覧のように全ての論文は右側にオープン査読のボックスがあり、この論文のステータスがどうなっているか説明されています。誰が査読者なのか、書かれています。名前が確認でき、面白いことに、査読者 2 は 3 人の査読者が一緒に査読をしているのだということが分かります。また、バージョンも確認できます。この論文には三つのバージョンがあり、各バージョンの査読ステータスを確認することができます。査読コメントを反映したり、ソフトウェアの論文であれば新しい事項について追記したりするなど、前のバージョンをリバイスしてアップデートすることもできます。新しいバージョンは独立して引用可能で、修正したいかどうか決めるのは著者です。著者がこの論文はこれで十分だと判断すれば、プロセスが終了します。

もう一つの重要なことは、論文のタイトルにバージョンナンバーと査読ステータスの詳細が含まれていることです。「査読待ち」から始まって、査読レポートの提出ごとにアップデートされます。このプロセス全体で出版を大幅にスピードアップすることができます。投稿から掲載までの平均時間は 7~8 日です。最初の

査読までの平均時間は 14 日、次の査読までの平均時間は 27 日、インデックス化までの平均時間は 34.5 日で、とても迅速です。

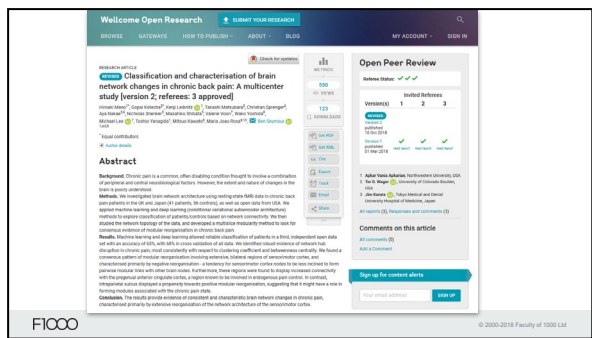
図 8 は、Ben Seymour の論文を示したものです。彼はこの論文についてこの後話をしてくれる予定です。オープン査読のボックスのリンクをクリックすると、査読レポートを見ることができます（図 9）。

「Continue reading（続きを読む）」をクリックすると、続きが表示されます。著者の回答を読むことができます。通常、「感謝します」「素晴らしいご意見です」といったことが言われますが、欠けているポイントを指摘したり、不同意を示したりして、査読者と議論することもあります。

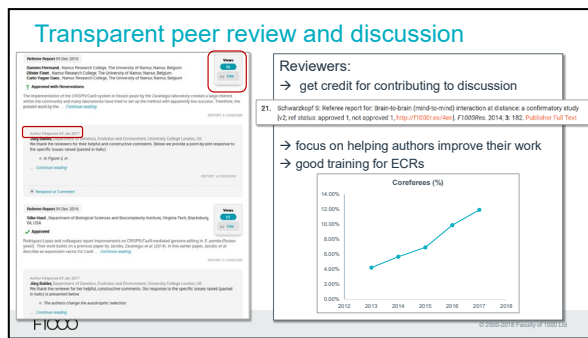
査読レポートも、論文とは独立して引用することができますので、査読者もクレジットを得ることができ、私たちは閲覧状況の追跡も行います。私たちの聞いたところでは、読者の多くはタイトルを見て、要旨を読んで、査読レポートを読んでから論文を読むかどうか決めています。査読者は議論に貢献すると評価されます。これは、著者ではなく査読者で引用された、査読レポートの引用の例です。

また、査読の性質も変わります。既に掲載されているので、論文をアクセプトするカリジェクトするかエディターが決められるよう手助けする必要はありません。査読者はただ単に、著者が論文の質を改善する手助けをします。それこそが本来査読が焦点を当てるべきことです。

また、これはキャリア初期の研究者にとってとても良いトレーニングになることが多いです。私たちが見



(図 8)



(図 9)

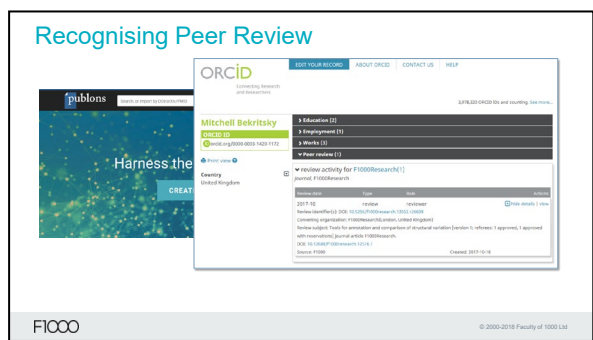
できた中で一つ興味深いのは、著者が複数いる、つまり2人以上で一緒に査読をして書かれる査読レポートの数が大幅に増加しているということです。大抵、ジュニアサイエンティストと一緒に査読レポートを書くことが多いですが、世界の反対側の科学者と協力してレポートを一緒に書くことも多く、これは興味深く重要なことです。

査読を行ってクレジットされるようにするツールは他にもあります。皆さんは Publons をご存じでしょうか（図 10）。これは研究者が全ての査読活動を把握できるようにするツールです。オープン査読であれば、査読レポートを読むことができますが、オープンでない査読でも査読活動を追跡することができます。また、自分の ORCID プロフィールを追跡することができます。ORCID の ID を持っていない研究者には、ID を取得するようお勧めします。簡単ですぐにできて、自分の研究活動を一つの場所にまとめることができ、とてもメリットがあるからです。

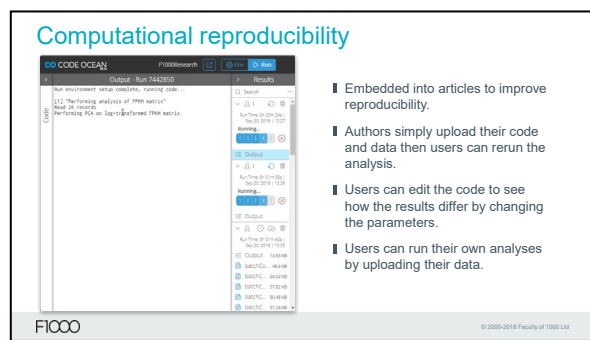
先ほど申し上げたように、私たちはデータの共有も

可能にします。私たちは、さまざまな事業者と協力して、論文の中にウィジェットを組み込み、読者や査読者がそのデータについて質問できるようにしました（図 11）。私たちは、例えば Shiny アプリ、R コード、他にもいろいろなツールを統合しています。また、計算再現性をサポートする統合も行っています。私たちは Code Ocean という組織と協力して、著者がコードとデータをアップロードし、その後、読者と査読者がそのコードを再実行できるようにしています（図 12）。コードは編集可能で、それによって分析がどう変化するか確認することができます。

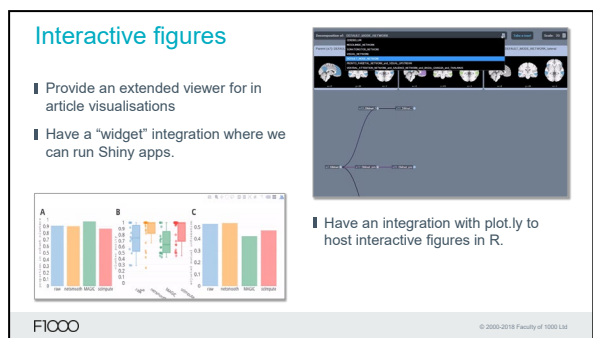
F1000Research は 6 年近く運営されており、私たちはこの期間中、このモデルの改善に取り組んできました。現在は、資金提供者などと共に、資金受給者に同様の出版モデルを提供することに取り組んでいます。これによりシステム全体を変え、研究者にインパクトファクターがなくても（そして私たちはインパクトファクターを導入しないと保証する方針です）、安全な方法で出版ができると安心させるのが目的です。また、資金提供者やさまざまな機関と協力して、システムの



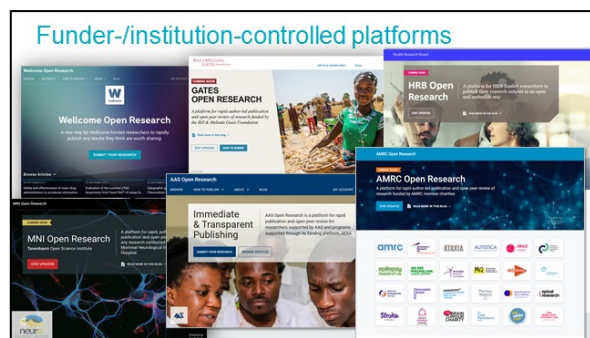
(図 10)



(図 12)



(図 11)



(図 13)

サポートにも取り組んでいます（図 13）。イギリスの Wellcome Trust は、私たちがサービスプロバイダを務める初のプラットフォーム、Wellcome Open Research を立ち上げました。これはこの組織の資金受給者向けのオプションのプラットフォームで、同じ出版モデルを採用しています。何が目的かというと、Wellcome がこう示すことです。「私たちは研究者であるあなたに資金を提供することにしました。そしてこれが、あなたが何でもすぐに共有できるようにするためのプラットフォームです」と。私たちは、ゲイツ財団、Irish Health Research Board など、多数のグループと同様のプラットフォームを立ち上げています。それから、欧州委員会も今年初めにフレームワークプログラム用の同様のプラットフォームの入札を行っています。

図 14 は、数週間前に「日本経済新聞」に掲載された、ゲイツオープンリサーチのプラットフォームに関する記事です。

私たちは、資金提供者のプラットフォームがそれぞれに競争し合い、インパクトファクターが生まれるような状況にしたいはありません。今は移行状態にあり

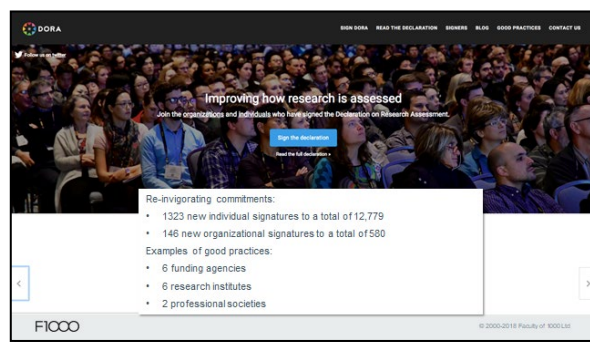
ます。私たちが最終的に到達すべき場所につながる道の一部なのです。最終的には、私たちは研究者が一元的に出版できるようになり、引用が問題でなくなり、一元的なプラットフォームが学術コミュニティで管理され、出版者はただこのようなプラットフォームで提供するサービスの質に基づいて著者の獲得競争を行うようになるべきだと考えています（図 15）。現在、私たちが質の指標として引用を使用することができないのであれば、そしてそれは当然行うべきでないのですが、そうであるなら依然として、資金提供やキャリアに関する決定が下せるようにするため、研究成果の品質レベル、重要性、潜在的な影響を評価するツールが必要なのは明白です。

先ほど申し上げたように、DORA は何年も前から研究者と研究が評価される方法の転換を促すことに取り組み、ますます多くの署名を集めており、再立ち上げ以降さらに多くの署名を得ていますが、DORA はグッドプラクティスの実例も集めており、私たちはそこから教訓を得てできるだけ真似したいと思っています（図 16）。

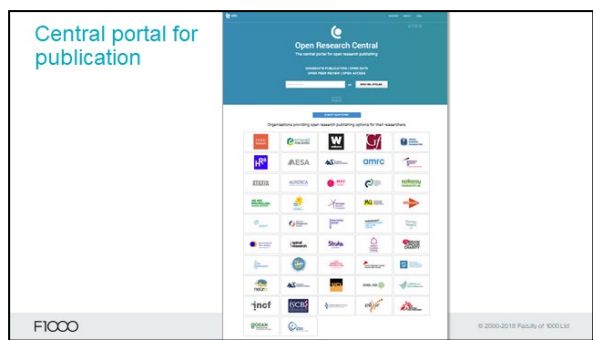
既にたくさんの、私たちが採録した利用可能なメ



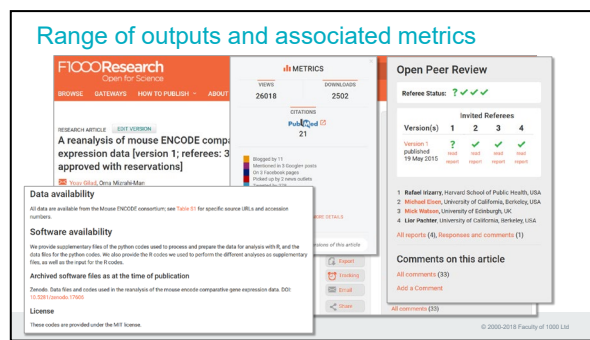
(図 14)



(図 16)



(図 15)



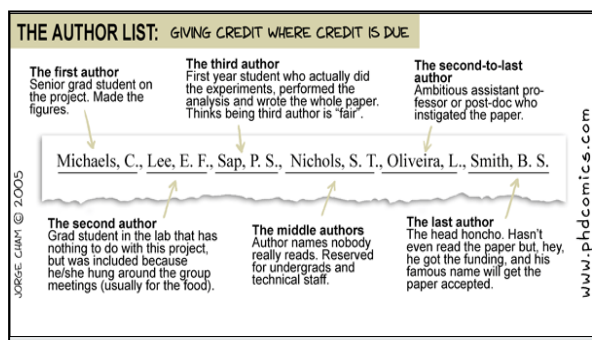
(図 17)

リクスがあります（図 17）。例えば、引用は依然として関連性があり重要なので、私たちはそれを採録しています。ソーシャルメディア活動やオルトメトリクスも、何を測定するかによりますが、特定の状況では重要になり得ます。また、私たちはオープン査読で捉えることができる興味深い重要な要素があると考えています。誰が査読者か、査読者が何と言っているか分かるからです。査読者は、最初から最後まで論文を読んだ唯一の人物であることが多く、従って、査読者が論文についてどう考えているか捉えることができるのは重要です。また、私たちは論文内のデータやソフトウェアといった要素に関するメトリクスや、引用や引用の使用法も把握することができます。

図 18 は、著者をめぐるよく起こる状況、誰が何に貢献し、誰が最初で誰が最後になるべきかといったこと全てを説明したものです。

Nature や PLOS などの出版者は、共同で CRediT の開発に取り組み、著者が通常論文に対して行う 14 の貢献の種類のオントロロジーを導入しました（図 19）。

主要出版者の多くは著者にこの情報を提供すること

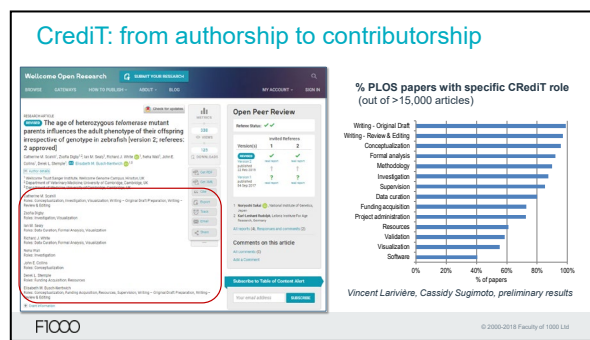


(図 18)

を求めています。図 20 はその例です。それぞれの著者が何に貢献し、それに従って個々のどのクレジットを得られるか、正確に確認できます。PLOS が 15,000 本以上の論文を対象に行った興味深い分析があります。この分析は、著者が貢献した活動の種類を示したものです。すると、例えば著者の貢献の約 80% はデータキュレーションであることが分かります。このような著者は大抵適切にクレジットされないで、優れたデータキュレーターを探したければ、その貢献を行った主要著者が誰なのか、知る必要があります。

F1000Prime

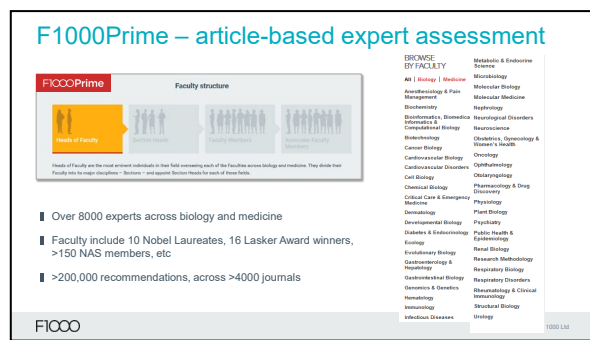
私たちは、品質と重要性を評価するツールも必要です。F1000 が最初に設立されてこのような名称になった理由は、私たちが F1000Prime というサービスを行っていたからです。これは、生物学と医学の分野で 8,000 人以上のメンバーを持つ大規模なバーチャルの専門家 faculty です（図 21）。何をしているのかというと、自身の研究の一環として文献を読む中で、特に興味深い論文や重要な論文を特定し、その論文の簡単な



(図 20)

Term	Definition	CRediT
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims.	
Methodology	Development or design of methodology; creation of models.	
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.	
Validation	Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.	
Formal Analysis	Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.	
Investigation	Conducting a research and investigation process, specifically performing the experiments, or data-evidence collection.	
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.	
Data Curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.	
Writing – Original Draft	Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).	
Writing – Review & Editing	Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.	
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.	
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.	
Project Admin	Management and coordination responsibility for the research activity planning and execution.	
Funding Acquisition	Acquisition of the financial support for the project leading to this publication.	https://forum.casrai.org/c/standards

(図 19)



(図 21)

推薦文を書くということです。面白いのは、彼らが推薦する論文の大半は、インパクトファクターがトップレベルのジャーナルに掲載されたものではないということです。このことは、たくさん的高品質な研究が至る所で発表されていることを示しています。これはグローバルな faculty で、とても高名な専門家がいて、この例のように特に日本に多く存在しています（図 22）。

図 23 は、推薦文を示したもので、推薦される論文と、その推薦内容が読めます。実際、論文はさまざまな専門家から多数の推薦を受けています。

質の指標

質の指標はさまざまなものがたくさんあります（図 24）。私たちは例えば、さまざまなグループとバッジ作りに取り組んでいます。さまざまなプレプリントやジャーナルでどのようなチェックが行われたかを知るのには難しいことが多いです。私たちは一連のバッジを作って、例えば盗作、倫理といった項目のチェックのレベルを管理しています。また、PreLights や PreReview のような、プレプリントなどを対象にした

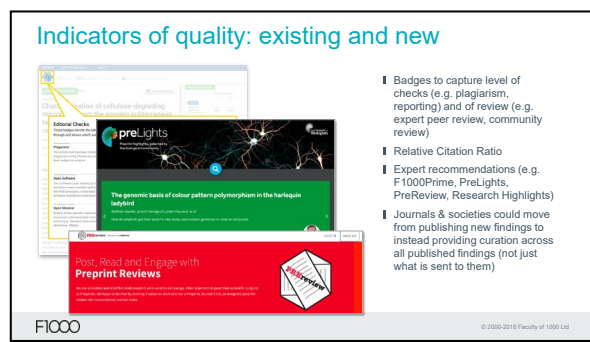
専門家の推薦ツールも多数あります。実際、ジャーナルの役割も変わり得るもので、『Nature』や『Cell』、『Science』といったジャーナルも、研究成果を出版する代わりに、一元的に発表された論文のキュレーションや強調することに焦点を当て、ジャーナルが特に重要だと思うもの、新規性と重要性に関する自らの基準を満たしたものを特定していくことができます。

まとめ

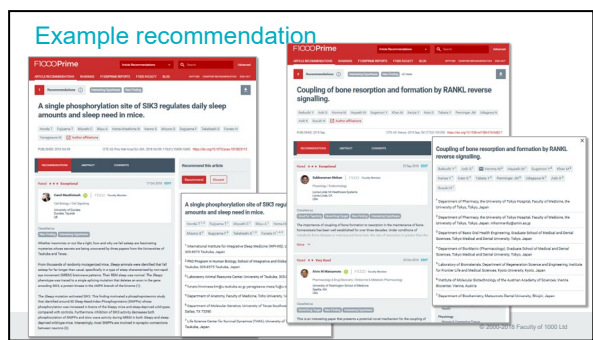
まとめると、オープンサイエンスへ向けた変化に伴う課題は、技術的な問題ではなく、文化的な問題だということです（図 25）。そのための鍵は、報酬とインセンティブのシステムを変えることです。私たちはもはやジャーナルを必要としていません。既に新しいモデルが利用可能になっていて、十分にテストされており、今では多くの資金提供者や政府が活用しています。私たちはこのようなアプローチへの世界的な転換の始まりを目にしています。こういったアプローチの幾つかを模索するため、日本の同僚の皆さんやさまざまなグループと取り組んでいければ幸いです。



(図 22)



(図 24)



(図 23)



(図 25)

第 2 回 SPARC Japan セミナー2018

「オープンサイエンス時代のクオリティコントロールを見通す」

学術コミュニケーションのエコシステムの今後 ～arXiv の現状から考える～

武田 英明

(国立情報学研究所)

講演要旨



インターネットの利用とオープンアクセス活動の浸透により、学術コミュニケーションのエコシステムが大きく変わってきている。伝統的なジャーナル投稿掲載という方法以外にも様々な方法が生まれている。現在の学術コミュニケーションのエコシステムを概観し、その中での preprint server の立ち位置を確認する。特に代表的 preprint server である arXiv を取り上げ、現状と課題を確認する。



武田 英明

SPARC Japan運営委員会委員長。

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

本日はいつもの話と少し違って arXiv、プレプリントを中心に話をしたいと思います。なぜ私がプレプリントの話をするかというと、10月2日に arXiv の Member Advisory Board (MAB) というミーティングに

行ってきたからです(図1)。arXiv というのはメンバー制度になっていて、大学等がメンバーです。日本の場合は日本コンソーシアムをつくっていて、幾つかの大学が入っています。そのコンソーシアムの代表が毎回 MAB に出ています。コンソーシアムの代表は、本当は京都大学の引原隆士先生なのですが、代表が代わる時期で、コンソーシアム代表代理という形で私が MAB に参加してきました。

MAB は 1 年に 1 回行われます。Member Advisory Board と、もう一つ Scientific Advisory Board (SAB) というボードがあり、この二つで arXiv はマネジメントされています。MAB の参加者の多くは図書館で、一部は Jisc などの研究助成機関です。研究者代表の方もいるのですが、今回は欠席でした。場所はコーネル大

arXivのMember Advisory Board (MAB) Meetingに行ってきました。

- 10/2 @Olin Library, Cornell University
- この会議はarXivのメンバーのうち、コンソーシアムなどの代表者などを招集して、1年に1度開催。
- 参加者
 - 多くの図書館：
 - Los Alamos Library, TIB, U of California, Ohio State Univ, U of Amsterdam, U of Queensland, U of Sydney
 - 研究助成機関
 - Jisc, Simons Foundation
 - 研究者代表: professor at KTH (欠席)

(図1)

学です。私は初めてコーネル大学に行きましたが、きれいで広大なキャンパスでした。

arXiv の現状と課題

現在、arXiv には約 140 万論文があり、1 日に約 600 件、新しい投稿が来ます。1 秒に 7 論文がダウンロードされます。モデレーターという、論文のチェックをする人が 162 人います。プレプリントは論文査読はしません。ただし、分野ごとにモデレーターのグループがあって、そのモデレーターが、その分野の論文として適切かどうかを評価します。それに通れば 1~2 日で掲載されるようになっています。

最近の傾向を見ると、2012 年のころはサブミッションが年間 84,000 件でしたが、2017 年は約 120,000 件で、5 割増ぐらいのペースで増えていることが分かります (図 2)。ポイントは、この増えた量がどこから来ているかということです。

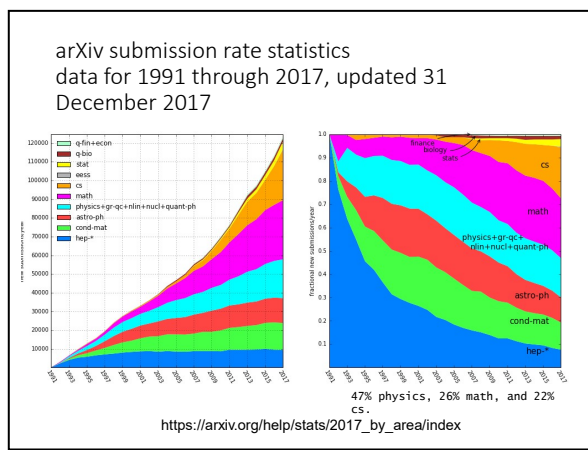
arXiv はできて 30 年たっていますが、図 3 の統計は 1991 年から始まっています。下から、青色が高エネルギー物理学、緑色が固体物理学、赤色が天体物理学です。これに水色を加えた下から四つが物理系です。さらに、紫が数学、オレンジがコンピュータサイエンスです。小さく見えるのが最近入ってきた統計など、広い意味での経済系です。

これを見て分かるように、元々の始まりである高エネルギー物理学はほとんど変わらない、減りもしなければ大して増えもしない状況です。それ以外の物理は

やや増えています。数学も比較的増えています。近年増加率が大きいのは、何といたってもコンピュータサイエンスです。2017 年は投稿の 22% がコンピュータサイエンスとなり、コンピュータサイエンスが顕著に増加しています。高エネルギー物理学は、相対的には arXiv の中のウエートとしては顕著に下がってきているということが分かります。とはいえ、物理全体で、5 割弱ぐらいは占めているという状況にあります。

これが分野としての分析です。これは図の下に記載されたサイトで手に入ります。

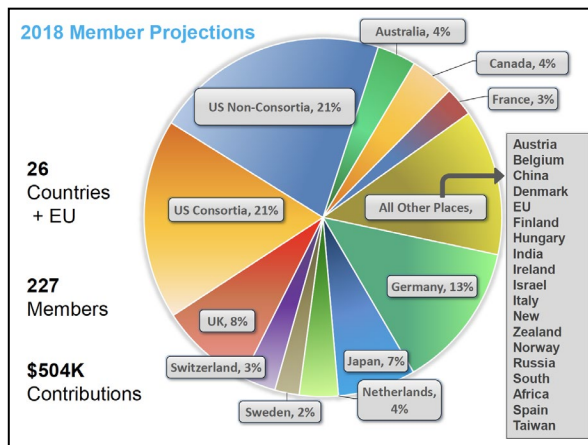
図 4 は、arXiv のメンバーはどここの国から来ているかという分布です。US Consortia が、アメリカの大学等がコンソーシアムを組んで加盟しているメンバーです。US Non-Consortia がそれ以外の独立で入っているメンバーです。それらを合わせてアメリカで約 40% です。そしてイギリスなど主要なヨーロッパの国があ



(図 3)



(図 2)



(図 4)

り、アメリカとヨーロッパ以外には、中国やオーストラリアがあります。日本も 7%入っていて、US、ヨーロッパ以外では最大となっています。

現在、200 強のメンバーが加盟しています。このメンバーはお金を払います。arXiv は多く投稿されている機関にメンバーになることを依頼しています。東京大学に所属する人が arXiv にたくさん投稿していれば、東京大学にメンバーになってほしいと依頼します。ただし、arXiv のメンバーになるのは強制ではありません。arXiv はご存じのように誰でも、所属があろうとなかろうと投稿できます。ただ、財政のためにこのような仕組みを取っています。

どこから投稿されているかという、メンバーの国と比例するように、アメリカが約 40%でヨーロッパは約 30%です（図 5）。中国はメンバー数に比べて投稿数が多く、3 番目となっています。これは課題です。

運営体制は、大きく、執行チームと運営チームと技術アドバイザーとなっています（図 6）。その他にアドバイザーのような形で、Paul Ginsparg がいます。彼は arXiv をつくった人で、今、コーネル大学のコンピュータサイエンス学科にいます。

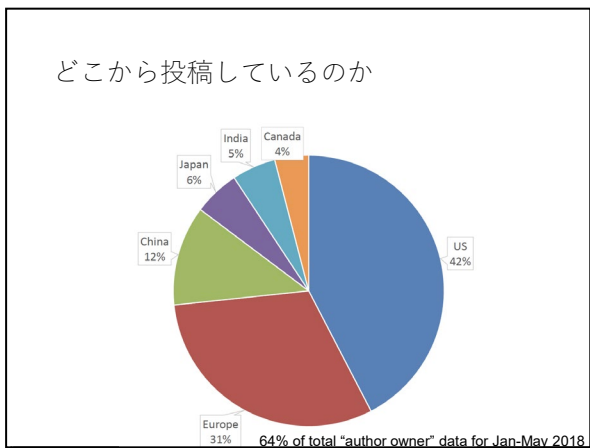
執行チームの Oya Riegerさんはライブラリアンです。今年までは arXiv はコーネル大学図書館が運営するという立場を取っています。だから、執行チームのトップの Oyaさんはライブラリアンなのです。

既にカレントアウェアネスでも報告が出たと思いますが、来年からこの体制が変わって、コンピュータサ

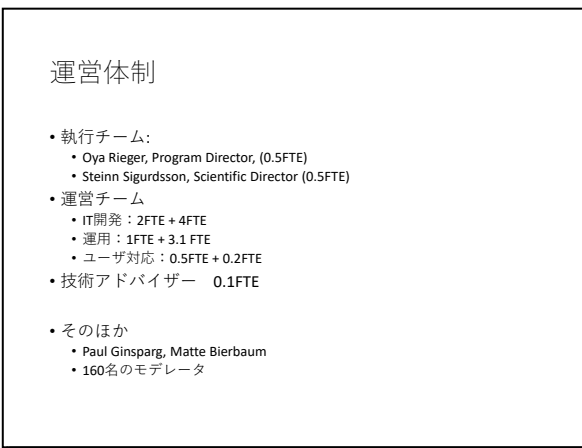
イエンス学科に移ります。われわれも心配したのですが、この運営体制そのものはほとんど変わらなくて、ただ、コーネル大学での責任を持つ部局が変わるということです。チームそのものが今、図書館の中にあるのですが、当面は物理的にも図書館の中にいるそうです。そういう意味では、ユーザの立場からですと、すぐ来年 1 月から何か変わるということはありません。

今まで、開発のことにに関しては図書館が責任を持つという体制で、執行チームと図書館が並行にあって、複雑だったのですが、コンピュータサイエンスに変わったら、今度は執行チームが開発チームを管理するという、むしろシンプルになると説明されました。Paul Ginsparg のアドバイスを受けて、より IT システムとしての改善を図りたいということが一つのモチベーションとなったそうです。

ビジネスモデルは重要なところで、これはサービスである以上、お金がないと回りません（図 7）。一体どうやって回しているか、大きく言うと三つのリソースからできていて、まずはコーネル大学自体資金を出しています。現金も出しているし、Oyaさんのようなコーネル大学所属の人に、人件費という形で出しています。それから、いわゆる財団系のもの、Simons Foundation 等が資金を日本円にすると 4,000 万円ぐらい出しています。あとは先ほど言ったメンバーから集めたお金が 5,000 万円ぐらいです。3 分の 1 ずつを負担しているようなイメージで動いているのが現在の仕組みです。総予算が約 1 億 5,000 万円です。



(図 5)



(図 6)

この金額が妥当かどうかというところは結構な問題なのです。arXiv を使ったことがある方は分かると思いますが、arXiv は実は結構、古くさいシステムです。それを今、ニュージェネレーションシステムに非常に時間をかけて移行しようとしています。歴史があるので、とても複雑なシステムになっています。LaTeX から入って、そのまま PDF が生成されるので便利なのですが、その分複雑なシステムになっていて、それがシステムをリニューアルするのにネックになっています。

今の話をまとめます。arXiv の分野としてコンピュータサイエンスが正式に入ったのが 2012 年ぐらいで、まだ 5 年ほどしかたっていません。それからさらに最近、統計や一部の経済系も入ってきています。先ほど表で示したように、arXiv は投稿数も増えているので、プレプリントサーバーを維持して、分野を順次増やしているという面においては arXiv は順調に成長しています。

また、arXiv がつくったプレプリントの文化およびビジネスモデルが、OA 時代を迎えて、分野を超えて高く評価されているということは大きいです。

図 8 が最近の動きです。arXiv は 1991 年に誕生しています。そこから 20 年以上、大した動きはなく、2013 年ごろから急激にプレプリントサーバーが他の分野で増えてきました。PeerJ のプレプリントが出てきたのが 2013 年で、2016 年以降、立て続けに出てきて、2016 年に bioRxiv、engrXiv、SocArXiv、PsyArXiv

が出てきています。2017 年には LawArXiv、AgriXiv、ChemRxiv など出てきています。

このうち、後ろに xiv と付いているものは、基本的には arXiv をリスペクトした、あるいはインスパイアされたそれぞれのサイトのページに書いてあります。arXiv のモデルをわれわれの分野でも使いたいと考えて立ち上げたと書いてあります。特に、赤いものが arXiv にインスパイアされたものです。

ここに違うもの、SSRN があります。これは約 2 年前にエルゼビアに買収されています。SSRN は元々、ソーシャルサイエンスのプレプリントサーバーでしたが、エルゼビアはこれをいろいろな意味で活用しようとしています。

その系列に当たるものが BioRN です。RN 系列のバイオ版をエルゼビアが立ち上げたということのようです。

また、同じ arXiv 系でも後ろに OSF と書いてあるのが、これは Open Science Framework で、Center for Open Science がつくっているプラットフォームに乗って、それを利用してサービスを提供しているプレプリントです。この五つは、マネジメントが違うだけでプラットフォームは同じものを使っています。それとは全く違う、インスパイアされたけれど、独自運営されているのは、bioRxiv、ChemRxiv です。これはアメリカとドイツとイギリスの化学学会が共同運営をしています。面白いのは、プラットフォームは figshare が提供するものを使っていると言っていることです。

ビジネスモデル

- コーネル大学: 170,000 USD + In kind
- Simons Foundation: 100,000 USD + 300,000 USD (マッチングファンド)
- メンバー機関: およそ 500,000USD
 - 機関あたり 1,000USD-4,400USD
- そのほか
- 計: 1,400,000 USD程度



(図 7)

台頭するpreprint サーバ

- 1991 arXiv: 物理学、数学、計算機科学、統計学、経済学
- 1993 SSRN: 社会科学 他 [Elsevier]
- 2013 PeerJ Preprints: バイオサイエンス [PeerJ]
- 2016 bioRxiv: バイオサイエンス [Cold Spring Harbor Laboratory]
- 2016 engrxiv: 工学 (OSF)
- 2016 socARXIV: 社会科学 (OSF)
- 2016 PysArXiv: 心理学 (OSF)
- 2017 BioRN: バイオサイエンス [Elsevier]
- 2017 LawArXiv: 法学 (OSF)
- 2017 AgriXiv: 農学 (OSF)
- 2018 ChemRxiv: 化学 [ACS, GDCh, RSC] (figshare)

(図 8)

2013 年、特に 2016 年以降にたくさんのプレプリントが生まれていて、ただ、多くは arXiv のつくった文化をむしろ自分の分野に取り入れたいという形で使われています。そういう意味で、arXiv が 30 年間行ってきた活動は大変高く評価されています。

ところが、この 30 年で学術コミュニケーションが変わってきました。オープンアクセスの動きが強くなってきましたが、30 年前はそのような動きはなかったのです。プレプリントサーバーの役割は変わってきています。arXiv が変わった環境の中できちんと対応できるかということが、arXiv という組織の課題となっています。

プレプリントサーバーの新しい役割

arXiv ができた当初、高エネルギー物理学では、新しい知見が発見されると、まずは関係する研究機関で情報を共有するという文化がありました。当初はファクスで送ったりするということがあったそうです。高エネルギー物理学の場合は、やっている研究所と大学が限られていて、世界中に拠点は数えられるほどなので、そういうことをやっていました。しかし、Paul Ginsparg が、それでは効率が上がらないから、一つサーバーを置いて、そのレポートを全部蓄積すれば楽だろうと考えて始めたのが arXiv です。大体的場合、そのレポートはいずれジャーナルにパブリッシュされるというものです。始まったころのプレプリントサーバーの役割は、比較的閉じたもので、あくまでも同業

者が知らせ合うというサービスでした。

ところが、2017 年、2018 年においては、いろいろな論文の発表の仕方があって、その一部としてプレプリントサーバーが位置付けられるようになりました。これは、当初のプレプリントサーバーの役割と違う役割を持たせるようになったということです。だから、プレプリントサーバー自体の機能もそのようにならなければいけないというのが、今の課題です。具体的には、他の publishing（出版・公開）の方法との役割分担、あるいは他の publishing との連携、さらには他のオープンアクセスへの積極的関与ということが求められています。

多様な publishing のやり方が生まれてきました（図 9）。では、具体的に何が違って来たかということです。元々のプレプリントサーバーの役割というのは、まず速報版を出して、それがジャーナルに投稿されるというものです。採録されても消されることはないの、それは著者版として残ります（図 10）。採録されないと、結果的にオリジナル版になります。つまり、ジャーナルには公開されないけれど、arXiv には残るのです。これが元々の arXiv の運用方法でした。それが逆に、初めからもう投稿しないつもりで出す人が出てきたのです。このプロセスに関係なく著者版を置きたいというだけのモチベーションで arXiv を使う人も出てきました。つまり、速報版、著者版、オリジナル版が同じ arXiv の中に共存しているということになります。

論文出版（公開）の多様性

- ・ジャーナル論文だけではなく多様な出版（公開）手段の発生
- ・プレプリントサーバ
- ・機関リポジトリ
- ・サブジェクトリポジトリ
- ・オンラインプロシーディングサーバ
- ・オーバーレイ出版

(図 9)

プレプリントサーバの役割

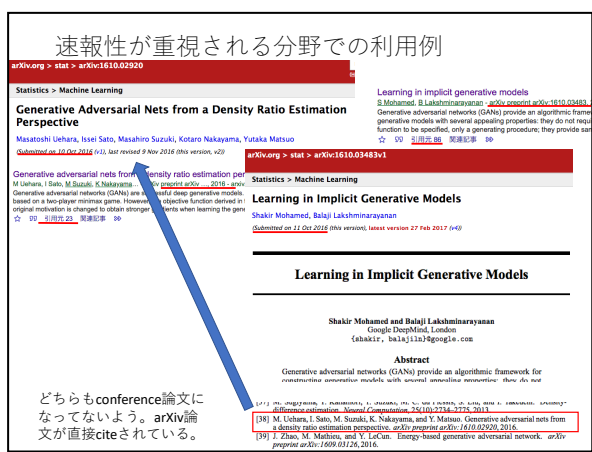
- ・速報版
 - ・発見などを素早く報告。その後、ジャーナル投稿、出版される。
 - ・著者版
 - ・OA版をプレプリントサーバにおく
 - ・出版版との共存の一つの手段
 - ・速報版もジャーナル論文が出版された後は著者版の役割
 - ・オリジナル版
 - ・プレプリントサーバの論文が原板となる論文
 - ・時間が勝負の論文
 - ・Citationされる
 - ・速報版もジャーナル論文が不採録の時はオリジナル版の役割
- ・これまで、プレプリントサーバの役割は速報版が第一義であったが、そのプロセスで生じる著者版、オリジナル版としての利用が出てきている。



(図 10)

特に arXiv は速報性が重視される分野で使われるようになってきました。図 11 は昔、北本朝展さんが紹介したものが面白かったのもう一回引っ張ってきました。この論文が 2016 年 10 月 10 日に arXiv に載せられると、別の arXiv の論文が引用するのです。その論文の出版日が 2016 年 10 月 11 日ということで、前日の投稿を引用しているのです。

調べたところ、この二つの論文ともジャーナルには採録されていないようなのです。でも、そんなことは関係ないのです。Google Scholar でどれだけ使われているかを見てみると、例えば右の論文は 86 件引用されています。左の論文も 23 件引用されています。ですから、別にジャーナルに採録されるかどうかは関係なく、arXiv にあるだけである種の目的は達成しているのです。



(図 11)

arXiv は OA に貢献しない？

もう一つは、ジャーナル、カンファレンス連携ということがあります。これは最終版とのリンクがあること、投稿時に連携することです。

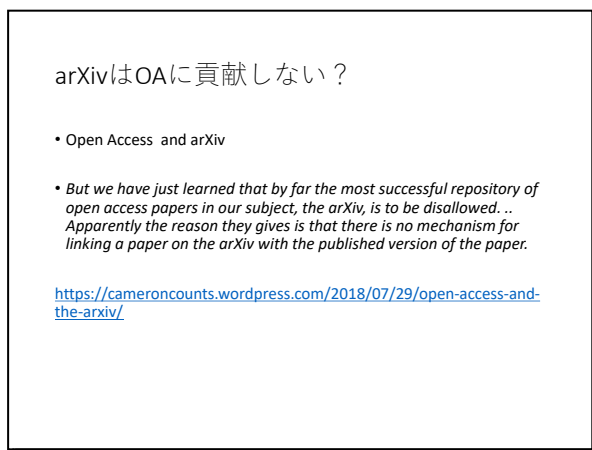
ある人が、arXiv は出版版とのリンクがないので、オープンアクセスの手段としてはカウントされないということを言っています (図 12)。実は、arXiv には明示的に出版版とリンクする仕組みがないのです。また、DOI もないのです。ここが問題ではないかということです。

後で坊農さんが説明する bioRxiv はきちんと DOI が付きます (図 13)。また、bioRxiv においてジャーナルに投稿するという仕組みが確立されています。従って、当然、投稿されると自動的に出版版とリンクが付くという仕組みを用意しています。

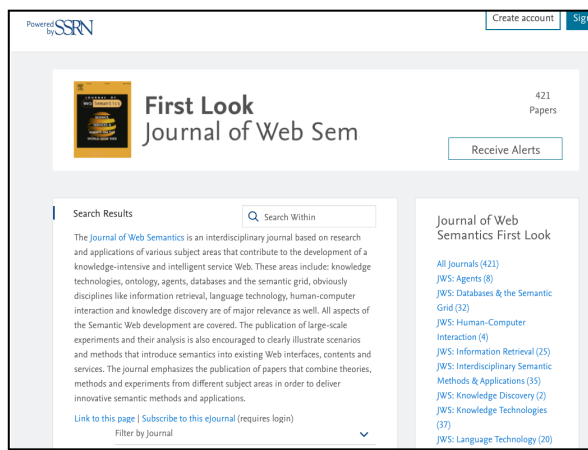
図 14 は私がエディターをしている『Journal of Web



(図 13)



(図 12)



(図 14)

Semantics』というエルゼビアが出しているジャーナルです。これは最近、投稿時は公開しないのだけれども、アクセプト時には著者版が SSRN に自動的に置かれるという機構を今度から導入すると言っています。今までのジャーナルはプライベートなサーバーに著者版を置いていたのですが、そうするとアクセスしづらいので、今までの 400 件の論文を全部含めて SSRN に置き換えたのです。

『Journal of Web Semantics』は、全くソーシャルサイエンスではなくてコンピュータサイエンスなのですが、なぜか SSRN に置くということです。これは恐らくエルゼビアがプレプリントサーバーを戦略的に使うための一つのトライアルです。こういうことが起きていて、これも著者版とのリンクです。

では、arXiv は今どうしているかということ、DOI 付与、メタデータの再設計に関してもまだめどが立って

いません。それはテクニカルな理由です。システムが直せないということです。メタデータがきちんとできないと言っていて、まだ見込みが立たないのです。

ジャーナル、カンファレンス連携に関しては、実験をしています。ACM と連携実験をしていて、まさに ACM のカンファレンスに投稿するときに、同時に arXiv に載せるという仕組みを今テスト中です。

今回の MAB にも ACM の人が来ていて、ACM および学会が arXiv とぜひ協力したいと。つまり、われわれはむしろ一緒に共存することが学会の使命だと思っているとはっきり言いました。ただ、arXiv 側がなかなかシステムの運用がままならないのが現状ということです。

また、arXiv Labs ということで、少しだけシステム改善で、サイテーションが下に出るような仕組みを実験的に導入しています。皆さんもぜひ試してください。

(図 15)

(図 17)

(図 16)

(図 18)

リファレンスとサイテーションが出るような仕組みができています（図 15）。

結論

なぜ研究者は論文を書くのか。それは、記録を残したい、人に知らしめたい、コミュニティで評価されたい、仕事を獲得したい・維持したい、これに尽きるのです。

17 世紀は、それを往復書簡で行っていました。それを書籍に書くと、人に知らせることができます。でも、評価されるのはなかなか難しいので、書評などがあります。

ジャーナルというのはある種の素晴らしいシステムで、記録、人に知らしめる、コミュニティに評価されたいというのをワンセットにして、非常によく回りました（図 16）。紀要は図 17 のようになります。電子ジャーナルは役割は紙版と同じですが、DOI など、より識別性が良くなりました（図 18）。

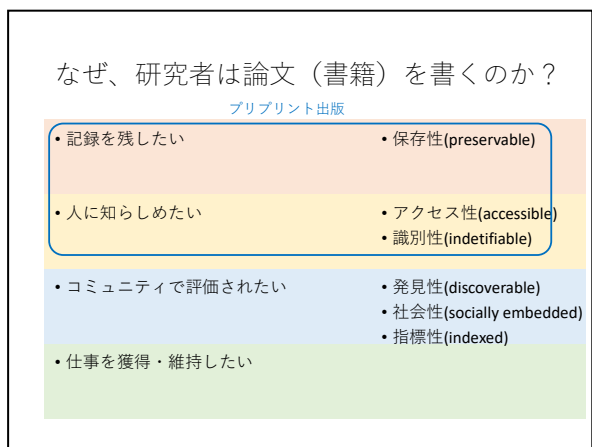
では、プレプリントではどうかというと、コミュニティで評価されたいというところまで入ってきません（図 19）。

そのときに、青い部分と赤い部分で役割分担をしようというのが今日の午前中の話でもあり、プレプリントの新しい役割です（図 20）。評価されるといった部分は別の機能に持たせればいいのか。そういうことが今起きているのがジャーナル連携です。

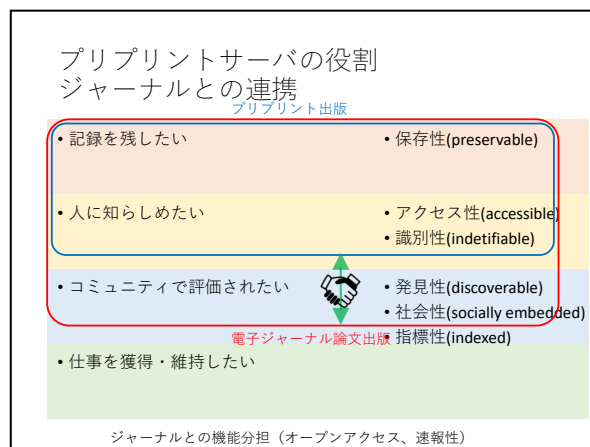
一方、直接オリジナル版でいいではないかという人

にとっては、むしろプレプリントサーバーの機能を強化してほしいというところがあります。機能を強化するということは、いろいろな指標性、アクセス性、識別性をプレプリントサーバー自体に持ってほしいという要求ということになります。

まとめると、プレプリントサーバーは求められる機能や役割が増えてきています。今回の MAB に出席して受けた印象は、arXiv はシステムはあまりモダンではなく、またマネジメントも比較的ゆっくりしているということです。ですので、現在の変化のスピードについていけるかという点を危惧しました。arXiv はサービス自身もサービスモデルも非常に期待されている一方、なかなか困難な段階になっていると感じました。



(図 19)



(図 20)

第2回 SPARC Japan セミナー2018

「オープンサイエンス時代のクオリティコントロールを見通す」

生命科学研究におけるプレプリント活用の現状

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)

講演要旨



プレプリントの活用が生命科学系においても急速に広まりつつある。プレプリントサーバーは、査読される前の結果だけでなく、これまで情報が共有されにくかった新型の計測機器によるデータの「癖」や、査読誌ではカバーしづらかった公共データベースやウェブツールに関する詳細な記述（ドキュメント）の置き場としても活用されている。2018年に入ってから、演者も生命科学系のプレプリントとしてよく使われるBioRxiv（バイオアーカイブ）に公共データベースからのメタ解析の研究論文を一本アップロードしたほか、別の共同研究の論文も査読誌に投稿する前にアップロードしてある。その現状について実例を交えながら紹介しつつ、いかにして情報の質を担保するかについて、演者の考えるところを議論したい。



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データベース (DDBJ) と連携して大規模塩基配列データ検索技術の開発とそれを活用した生命科学研究を行っている。京都大学博士 (理学)。

私の所属は情報・システム研究機構のライフサイエンス統合データベースセンターで、静岡県三島にある国立遺伝学研究所のDNA Data Bank of JapanのDNAデータを扱いやすくしたり、実際にそのデータを使った研究もしたりしています。

私はこの1年でプレプリントが身近になったので、これを機会に自分の1年をまとめてみたいと思います。

1~2年でオープンアクセスはどう変わったか

図1は、私が2年前にこのSPARC Japan セミナーで出した、生命科学研究者の立場から見てオープンアクセスはどのようなものかという雑感を書いたものです。この2年で変わったのは、bioRxiv というものが、新しいことを考えている人には知られるようになったと

いうことです。生命科学研究者は、『PLOS ONE』や『Scientific Reports』といったオープンアクセス誌のことは知るようになって、自分もそういうところに出たいと思うようにはなりました。しかし、グリーン

生命科学研究者の立場から見たOpen Access (OA)

・現状、ゴールド OA一択

・ すなわち「オープンアクセス」=ゴールドOA

+ BioRxiv

・グリーンOAという言葉自体ほとんど知られてない

- ・ IFの付いてない雑誌は回避される傾向
- ・ ピアレビューを経た論文のみ業績とみなす伝統

第1回 SPARC Japan セミナー2016 「オープンアクセスへの道」
坊農秀雅「生命科学研究分野における研究者の投稿先雑誌選択傾向とOAへの意味づけ」
スライド一部改変

(図1)

OA という言葉はほとんど知られていない、インパクトファクターの付いていない雑誌は回避される傾向がある、ピアレビューを経た論文のみ業績とみなす伝統があるという状況はあまり変わっていないのではないかと思います。

ちょうど昨年の今ぐらいに、『Science』という雑誌がプレプリントに関する特集を組んで、生命科学でもプレプリントがあり、論文が出されているということを紹介してくれました（図 2）。このスライドも去年出しています。

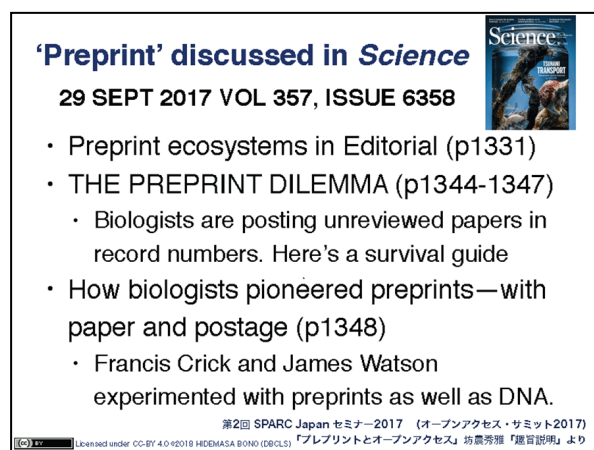
私が bioRxiv フリークになった理由

昨年、ここで 10 月最後の月曜日に SPARC Japan セミナーに出させていただいて、その週の途中から Cold Spring Harbor 研究所に出張で参りました（図 3）。私の研究分野はパイオインフォマティクスという分野

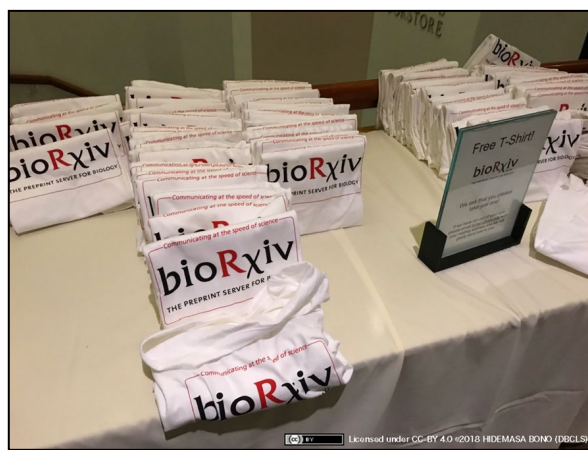
で、生命科学でも普段から実験するのではなく、データを相手にして、DNA のデータから何か新しいものを見つけたり、データベースをつくったりしています。その研究で有名な研究所です。DNA の二重らせんのモデルのブロンズ像が置いてあります。bioRxiv は Cold Spring Harbor Laboratory Press というこの出版会がやっているものです。

そこで学会があったのですが、その学会に行くところんなものが置いてあって、たくさん T シャツを配っていました（図 4）。今日はその T シャツを着ています。帰ってベッドで広げて写真も撮りました。自分のノートパソコンに bioRxiv のシールを貼っているぐらい、bioRxiv フリークになりました。

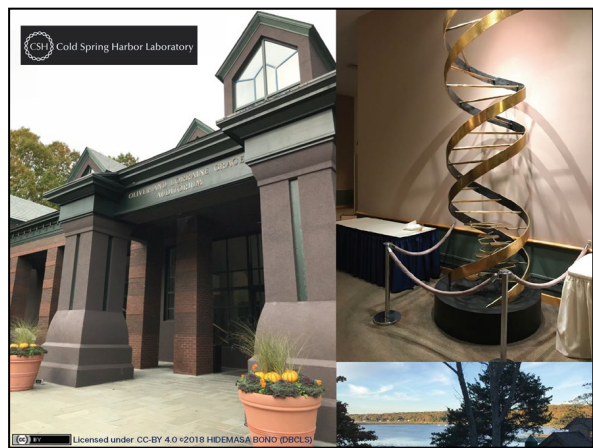
それは理由があって、皆さん、学会で発表するからです。口頭発表とポスター発表があり、口頭発表者はもちろんなのですが、ポスター発表者でも、全部自分



(図 2)



(図 4)



(図 3)



(図 5)

がやっている研究は bioRxiv に論文としてアップロードしているのです。昔だったら、「まだ論文投稿中で、あまり詳しくは話せないのだけれども」と言っていたのが、bioRxiv の ID、つまり DOI をみんな書いてあるのです。それにショックを受けて、「これは自分もやらなければいけない」と思いました。

「次の論文はプレプリントサーバー (bioRxiv) にまず出そう」という決意をして、その決意を証拠に残すために Twitter でもその日につぶやきました。11 月 4 日という、去年の今ぐらいですが、証拠が残っています (図 5)。

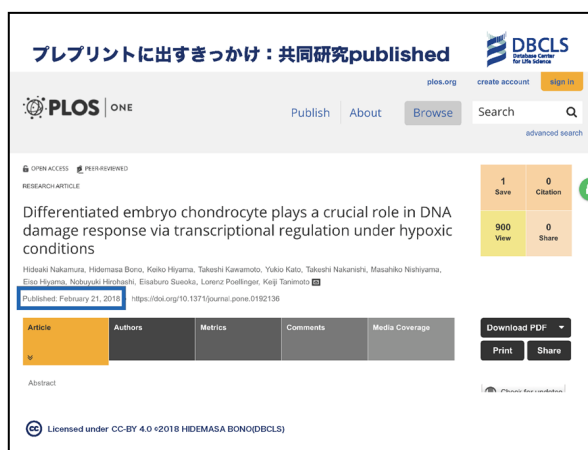
その雑感をさらに Twitter に書いて「月曜日の SPARC Japan セミナー『プレプリントとオープンアクセス』に始まって、水曜日から土曜日まで bioRxiv をホストしている Cold Spring Harbor Laboratory での Genome Informatics Conference。すごくプレプリントな 1 週間だったなあ」と言っていたぐらい、その 1 週間でプレプリントをやりたい、出したいという人になりました。

bioRxiv へのサブミット

帰ってすぐやれたらよかったのですが、いろいろやっていて 2 月になって、出す機会がついにめぐってきました。図 6 は私が出したプレプリントのウェブサイトのスクリーンショットです。「Meta-analysis of hypoxic transcriptomes from public databases (公共データベースの中の低酸素刺激による遺伝子発現データのメタ解

析)」という題名の論文をプレプリントサーバーにアップしました。これはなぜ出したかという、きっかけがありまして、図 7 も実は『PLOS ONE』に出したのですが、私の共同研究の論文が、2018 年 2 月 21 日に出たのです。私自身がデータベースから検索して見つけてきたことをきっかけに、共同研究者が実験をして発見があった、その報告を『PLOS ONE』に出したのです。その私が見つけてきたというのが、ただ単にデータベースを検索したわけではなく、データベースからいろいろデータを集めてきて、バイオインフォマティクスの計算をして、キュレーションをした結果、見つけられました。

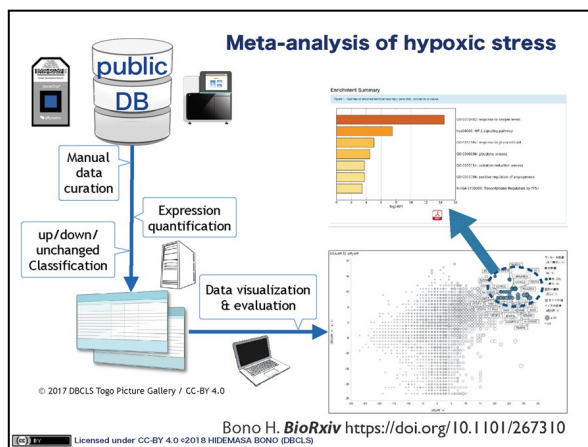
その結果も、ただ単にこれを見つければいいという一個だけの遺伝子ではなく、たくさんある中から見つけてくださいと共同研究者に遺伝子リストを渡してありました。共同研究者はそれを長らく 2 年ぐらい見て



(図 7)



(図 6)



(図 8)

いて、その結果見つけてきて論文になったのです。

私自身がやった手法を、本論文には詳しくは書けなかったもので、それをどこかに残したいからプレプリントに書こうということで、ずっと前から書いてあったのですが、この論文が出版されたので、それに合わせて bioRxiv に投稿しました。これは日にちが大事なのですが、先ほどの論文は 2 月 21 日にパブリッシュされ、私はその日のうちに bioRxiv にサブミットしました。

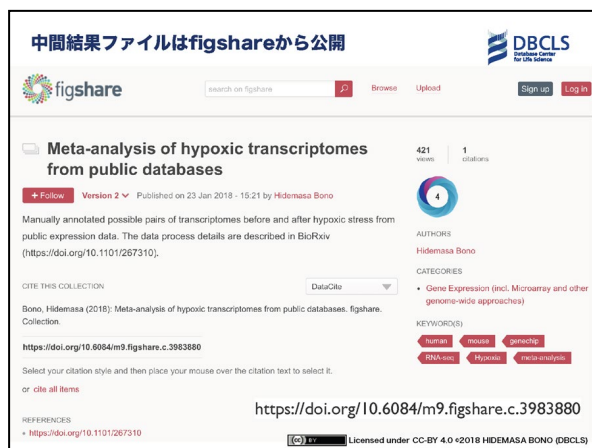
2 月 23 日にこれは間違いなくオリジナルの論文だと認めていただいて、DOI もめでたく付いて、公開されることになりました。この日は静岡県民にとっては大事な日で、富士山（223）の日です。富士山の日にめでたく出せたということです。

この研究は実際に何をやっているのかというのを少しだけ紹介させていただくと、公共データベースを使

っていて、これは遺伝研にある DNA のデータベースなのですが、その DNA のデータベースを使って、人間が酸素がうすいところに行ったらどんなことになるかを調べました。これは人間の実験で、あまりやると死んでしまうので、できないのです。普通の実験で使う酸素濃度はとてもうすいです。私たちがよく使うのは 1%の酸素濃度です。普段皆さんは 21%のところに住んでいますが、1%の酸素濃度になったときにどういうふうに変化するかを見るために、いろいろな人がおこなった実験からデータを取り出して比較しました。

DNA の配列そのものではなく、タンパク質になる前に出てくるメッセンジャーRNA の数がどれぐらい変わるかを調べました（図 8）。

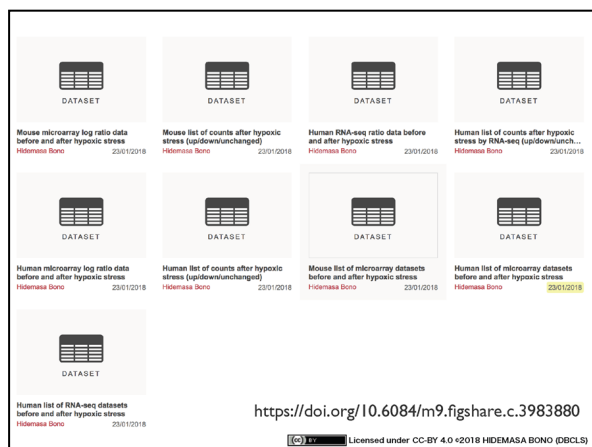
そういうことをおこなったデータがたくさんデータベースに入っているのを、それを計算したデータのフ



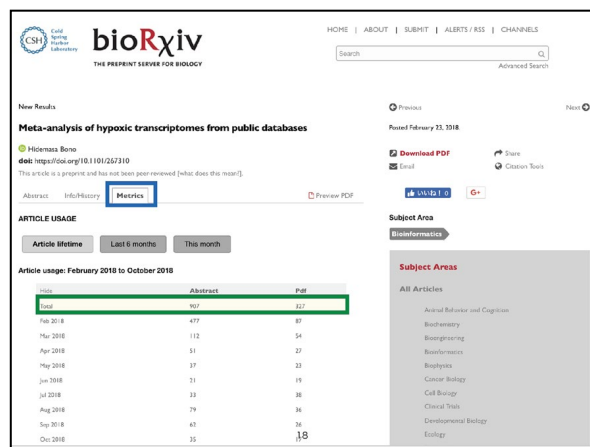
(図 9)



(図 11)



(図 10)



(図 12)

ファイルをつくりました。最終的には、論文にするときは可視化するのですが、実際に大事なものはこのような中間データです（図 9）。このあたりのデータは専門家が見るとよく分かるので、中間データを皆さんでシェアしたい、このデータは再利用しがいがあるということで、bioRxiv に出しました。ただ単に bioRxiv に出すだけだったら普通の成果なのですが、このようなテーブル形式のデータをサブミットしたいということで、テーブル形式の Excel のデータは figshare にアップしました。

figshare にアップすると、同じようにデータテーブルに DOI が付与されます（図 10）。ここからダウンロードできるようにして、この figshare の DOI を先ほどの論文にたくさん書いて登録しました。

データテーブルの中身は図 11 のように、遺伝子の ID と数字がたくさん書いてあります。

bioRxiv の「Metrics」のタブを選択すると、何回見られたかが分かります（図 12）。これは 2 月にアップしたばかりですが、おかげさまでトータルで 900 回ぐらい見られています。論文全文の PDF がダウンロードされた回数も月ごとに出て、よく分かります。これも論文をプレプリントにアップした人間としては励みになります。

実は、こういうことがあるよと私の共同研究者に言うと、共同研究者は私より前に bioRxiv にアップしていたのです。その共同研究者から、一緒に論文をアップしようということで、2018 年 7 月にアップして、3

カ月ぐらいでこれぐらい見られていますということが分かりました（図 13）。自分がメインで出した論文を 1 本、共同研究でもう 1 本、bioRxiv にこの 1 年で出すことができました。

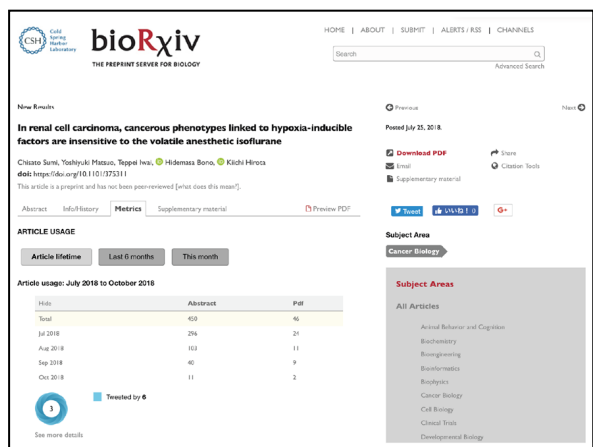
生命科学研究者は、アメリカ国立衛生研究所（NIH）の国立生物工学情報センター（NCBI）が運営している PubMed で論文を検索して、そこに載っていたら、みんなが成果と認めるという暗黙の了解があるのです。PubMed ではまだこのプレプリントは検索できないのですが、Europe PMC では検索できます（図 14）。

Europe PMC をご存じない方もいるかもしれませんが、ヨーロッパでも PubMed 的なアクティビティをしていて、PMC は PubMed Central の略で、アメリカの PubMed の全文版が PMC と呼ばれて、利用可能になっています。PubMed はご存じのとおり抄録、アブストラクトだけなのですが、PMC は全文が見られるコレクションです。これも数が増えています。

基本的には同じデータがあるのですが、Europe PMC のインターフェースでは、例えば私の名前で検索すると、bioRxiv に登録されたようなプレプリントも一緒に検索できます。このようにプレプリントと分かるように書いてあって、PubMed ID は当然ないのですが、検索できるという状況になっています。

プレプリントのメリット・デメリット

出してみても分かったプレプリントのメリット・デメリットについてお話しします。メリットは、すぐ公開



(図 13)



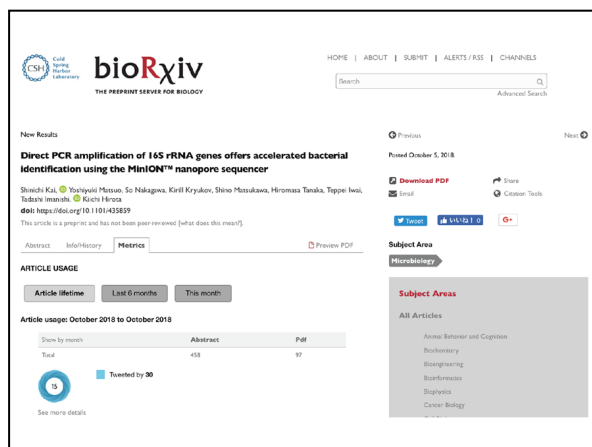
(図 14)

できる、DOI で引用可能である、現状は article processing charge (APC) 無料、査読誌への投稿ができる、Europe PMC から検索可能という点です。特に私がメリットだと思ったのは、DOI で引用可能なために、研究費申請のときにリファレンスとして引用することができるとことです。また、私が勧めた友達が良かったと言っていたのは、先に論文を出せたということです。生命科学者の中では、論文を投稿している間にレビュアーがそれを見て、それをまねして実験を始めて、出し抜くことがあるといわれています。それが阻止できる、先取性が担保できるのがいいと思っています。

現状で感じているデメリットは、この辺はよく理解されていないから起きていることなのですが、一度そのように公開すると別のところに論文が出せないかもしれないという恐怖です。ほとんどの場合は時代が追い付いてきて、そんなことはなくて、多くの雑誌に論文を出すことはできるという意識が広まると思います。

また、今までやってきた査読論文に出すスタイルを変えたくないとか、査読論文しか評価されない現実があるとか、そういう後ろ向きなことですが、このあたりもできれば変えていけたらと思っています。

私の共同研究者は、果敢にも今月も、とても注目されるようなプレプリントを出しています(図 15)。これは査読論文を出す前にこういうことを出しているのですが、今月だけで 500 を超えるぐらい見られていて、こういうものをうまく使うと注目される、という形に



(図 15)

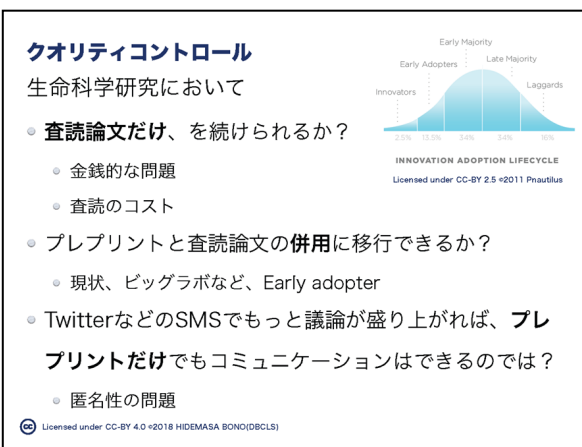
できるのではないかと考えています。

クオリティコントロール

最後にクオリティコントロールの話をしします。われわれの生命科学の研究において、今までやってきた査読論文だけを続けられるかということがあります。査読にはコストがかかります。午前中に、『Scientific Reports』がだいぶ遅くなっているという話がありましたが、それは私たちも感じています。査読していただくということだけでは無理ではないかという実感があります。

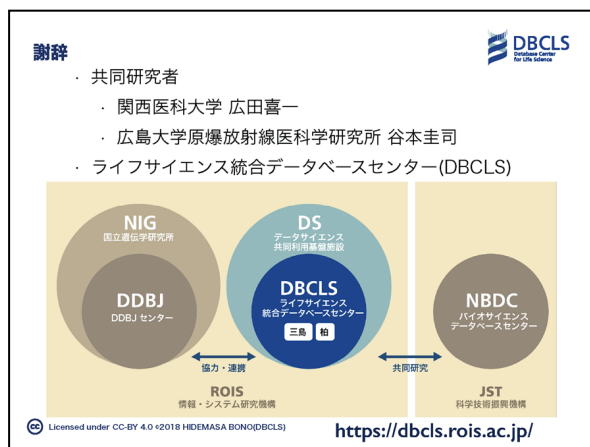
次の段階として、プレプリントと査読論文の併用に移行できないかと考えていますが、プレプリントを出すというのは、図 16 の分布で言うところの Early Adopters ぐらいしかできていないので、まだ様子見のレベルではないかと思います。もっと議論が盛り上がっていけば、プレプリントだけでもっとコミュニケーションができるのではないかと思います。

海外は分からないですが、日本の Twitter をする人は、研究者と名乗っては言っているものの、匿名でやりたい人が多くて、議論は盛り上がっているようなのですが、誰か分からないということが多く、そのあたりが邪魔するのだろうかと考えています。ぜひ実名でこういう議論ができるようになればいいのですが、実名でやってしまうと、自分のプロモーションなどに影響するのではないかと考えている人が多いのではないかと思います。



(図 16)

この研究は、関西医科大学の広田喜一さんと、広島大学の谷本圭司さんという方との共同研究で行ったものです（図 17）。普段こういう研究をする機会を与えていただいているのは、ライフサイエンス統合データベースセンター（DBCLS）です。DBCLS は図の真ん中の組織で、今回の SPARC Japan セミナー企画ワーキンググループメンバーであるバイオサイエンスデータベースセンター（NBDC）の八塚さんとは共同研究を普段させていただいています。この国立情報学研究所も情報・システム研究機構（ROIS）の中の研究所であるということを最後に紹介して終わりにしたいと思います。



(図 17)

第2回 SPARC Japan セミナー2018

「オープンサイエンス時代のクオリティコントロールを見通す」

パネルディスカッション



林 和弘 (科学技術・学術政策研究所)

Rebecca Lawrence (F1000)

武田 英明 (国立情報学研究所)

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)

栗山 正光 (首都大学東京)

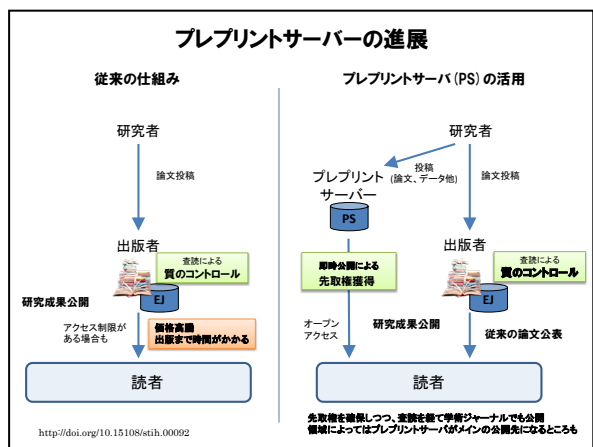
●林 パネルディスカッションに入る前に、今回もさまざまなセクターからご参加いただいているので、そもそも今日お話ししている根本のモデルについて、ご紹介したいと思います。

普段、研究者は査読付きジャーナルに出すのですが、ジャーナルの価格が高騰している、質のコントロールが難しい、レビューに時間がかかる、レビュー者を見つけるのも大変だしその中身に問題があるなど、いろいろと課題がある中で、プレプリントサーバーに出して先取権をまず確保してオープンに出そうという動きがあります(図1)。高エネルギー物理学などはそれとは別に査読付きジャーナルにも出して、実績は改めて獲得しますが、査読の時間を待ってられないような、研究サイクルが短い分野ではプレプリントサーバ

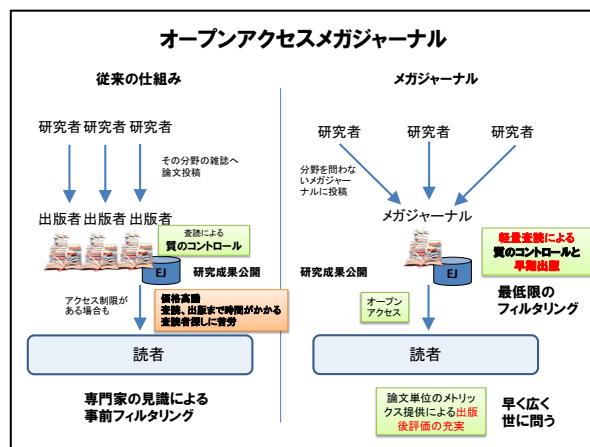
ーをより重視する動きが加速化している状況にあります。そこから、誰がプレプリントの質保証をするのかという今日のテーマとつながります。

それとは別に、オープンアクセスメガジャーナルも一つのゲームチェンジのモデルです(図2)。それぞれの分野のジャーナルに出すのではなく、一つにまとめて出して、Benさんの紹介のとおり最低限のフィルタリングをして、あとは世の中にオープンにその評価を問いましょうというものです。このピアレビューの質とトラディショナルなジャーナルのピアレビューの質はどこまで一緒に、どこまで違うのかという議論もあると思います。

そして、今日 Rebecca さんにお話しいただいたメイントピックの一つは、研究助成機関がサーバーを持つ



(図1)



(図2)

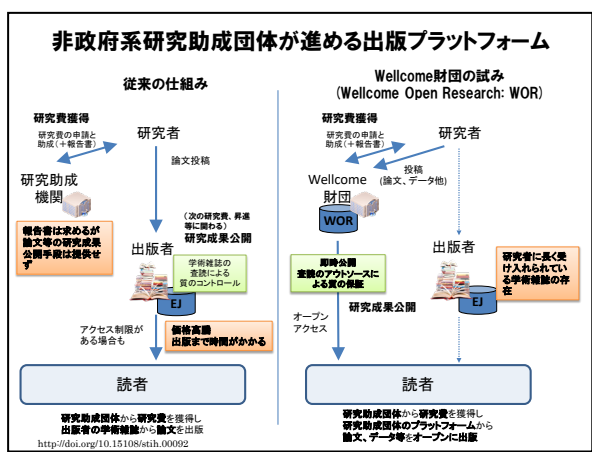
ということでした（図 3）。今までは助成機関は報告書をもらうぐらいで、成果はあくまで既存の出版社から出していたのですが、Wellcome Open Research 等のサーバーでまずプレプリント相当で出してしまっ、即時公開でタイムスタンプを付けます。その後、Faculty of 1000 でいわゆる査読のアウトソースをして、クオリティコントロールをして、通った論文はPubMed 等にインデックスされます。

この仕掛けがどのくらい通用するのか、Ben さんの言説で言えば、長く研究者に受け入れられている学術雑誌の存在は無視できない、インパクトファクターは無視できない、坊間さんからもそういうお話があったかと思ひます。

これに関して、実は私はバックデータの一つ持っていて、ちょうどいい機会なのでご紹介しします（図 4）。2013 年に、「あなたが一番良い成果を出したら、どの

ジャーナルに出したいですか」と 8 分野に分けて聞きました。要は『Nature』と『Science』信仰はどれぐらいあるかというのを知りたかったのです。私は作業仮説として、欧米の国際学会のトップジャーナルが強い場合もあるのではないかと考えていて、実際に分野によっては『Nature』『Science』よりも出したいジャーナルが各分野にあることが分かり、ある意味ほっとした面もありました。一方で工学や地球・環境は『Nature』『Science』がいいというのが定量的にも出ているという状況です。日本だけではなく世界中の研究者に聞いた結果です。実はこのようなデータを基に Nature Index がつくられています。いずれにせよ、研究者には出したいジャーナルというものが存在しているという状況下でプレプリントサーバーの扱いをどうするかという議論があります。

それから、もはやジャーナルである必要はないという Rebecca さんのお話がありました。一番進んでいるというか、一つの試みとして Registered Report というものがあります（図 5）。そもそも実験する前に論文を書いてしまうというか、実験計画を公表してしまうわけです。そこで査読をして、その後、実験をしてデータを取ったら、その結果が作業仮説に対してポジティブだろうが、ネガティブだろうが出版してしまうという仕組みです。このようなメディアも既に一部の分野で、臨床系が中心と理解しているのですが、動いています。ある意味、ジャーナルというシステムを根本的に壊す仕掛けなのかもしれません。ポジティブに言



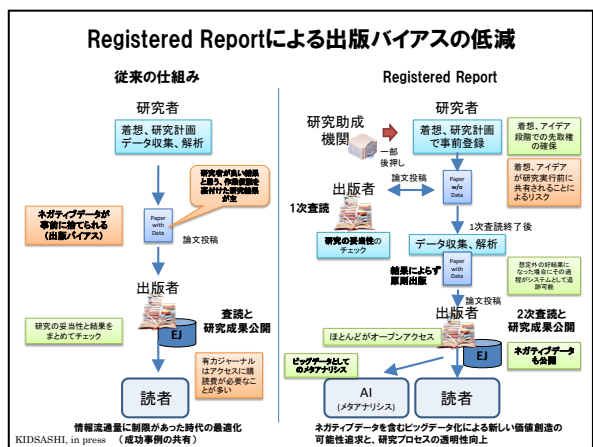
Which journal do you wish to submit with your best results?

分野	英語	日本語	英語	日本語	英語	日本語	英語	日本語	英語	日本語
1	SCIENCE	44	NATURE	1162	JAM CHEM SOC	264	SCIENCE	149	1	
2	ASTRONOMY & ASTROPHYSICS	43	SCIENCE	1124	AMERICAN CHEMICAL SOCIETY	230	NATURE	146	2	
3	NATURE	41	P NATL ACAD SCI USA	868	SCIENCE	221	NAT GEOGRAPHIC	125	3	
4	NATURE	40	CELL	734	NAT CHEM	215	GEOGRAPHY RES LETT	87	4	
5	NATURE	35	NAT CELL BIOL	497	NATURE	197	EARTH PLANET SCI LETT	77	5	
6	NON NOT R ASTRON SOC	35	NAT CELL BIOL	331	CHEM COMMUN	148	GEOLOGY	73	6	
7	ICARUS	19	EMBO J	319	J ORG CHEM	116	J GEOPHYS RES	67	7	
8	J GEOPHYS RES	13	J ORG CHEM	319	CHEMISER J	109	P NATL ACAD SCI USA	61	8	
9	GEOGRAPHY RES LETT	12	NAT MED	309	ORG LETT	107	J CLIMATE	33	9	
10	NAT PHYS	10	NAT GENET	264	P NATL ACAD SCI USA	85	GEOCHIM COSMOCHEM ACTA	31	10	
総計	375	総計	10078	総計	2081	総計	1009			
回答数	54	回答数	1521	回答数	380	回答数	216			

分野	英語	日本語	英語	日本語	英語	日本語	英語	日本語	英語	日本語
1	SCIENCE	93	NAT MATER	142	NEW ENGL J MED	719	PHYS REV LETT	267	1	
2	NATURE	87	ADV MATER	112	LANCET	638	SCIENCE	261	2	
3	PHYS REV LETT	40	NATURE	110	JAMA JAMA ASSOC	445	NAT PHYS	201	3	
4	P NATL ACAD SCI USA	40	SCIENCE	106	BRIT MED J	299	NATURE	192	4	
5	PHYS REV LETT	38	NANO LETT	83	PLoS MED	256	PHYS REV B	110	5	
6	APPL PHYS LETT	36	NAT NANOTECHNOL	74	CIRCULATION	187	APPL PHYS LETT	94	6	
7	NAT NANOTECHNOL	31	PHYS REV LETT	62	J CLIN ONCOL	167	P NATL ACAD SCI USA	71	7	
8	NANO LETT	31	ADV FUNCT MATER	60	LANCET ONCOL	145	NAT MATER	70	8	
9	NAT MATER	30	APPL PHYS LETT	55	ANN INTERN MED	143	NAT PHOTONICS	62	9	
10	NAT PHYS	24	JAM CHEM SOC	50	BLOOD	121	NEW J PHYS	60	10	
総計	1168	総計	1521	総計	121	総計	2206			
回答数	169	回答数	195	回答数	7205	回答数	201			

研究・技術計画学会 第28回年次学術大会 (2013) 2112

(図 4)



(図 5)

えば、いわゆる出版バイアスを経ることなく、とにかく研究者がある作業仮説に基づいて行った研究に関する成果をニュートラルに、オープンに公開しましょうという動きもあるということです。

今までご紹介した事例は、紙メディアでは絶対にできなかったのです。ウェブ上に知識を置いて皆さんに共有する、情報が誰にでも届く時代になったときに、では、そのコンテンツの質の保証は誰がどうするのか、あるいは伝達のコストはどう賄うのか、新しいモデルが見通せたときに、それに向かってどのように変わっていくべきかをこれから議論したいということで補足させていただきました。

それでは、ディスカッションに入りたいと思います。最初に、Rebecca さんに後半の武田先生と坊農先生の講演について感想を伺いたいと思います。

●Lawrence プレプリントに関するお二人のお話はいずれも素晴らしいものでした。arXiv は特にこの道を主導しています。物理学の分野では長年にわたって多くのプレプリントが行われているのに、依然として成長しているというのは興味深いですね。実際、どんどん多くの分野がプレプリントを行うようになってきています。実際には、生物学分野の bioRxiv は依然として出版される全論文のほんの一部しかカバーしていませんが、これも大きく成長しています。プレプリントは前進に向けた大きな転換となっていますが、プレプリントとジャーナルが別々にあるというのは理にかないません。プレプリントの上に査読やキュレーションなど、その他みんなが希望するさまざまなタイプのチェックを重ねることが先に進むためのより良い道です。

●林 ありがとうございます。

次に、ありがたいことに、ご質問も多く頂いています。びっくりするくらいあります。

Rebecca さんに対する質問です。リジェクトされたペーパーを見られるようにした場合、マイナーなミス

を修正したペーパーを他人が出版するという、スクーピングが発生することもあります。このマナー違反の問題をどのように扱いますか。

●Lawrence 覚えておかなければいけないのは、F1000Research では「リジェクト」（または「非承認」）されても掲載されるということです。既に掲載されたものをスクーピングすることはできません。問題が見つかったら、それに基づいて修正できます。科学とはそういうものです。何かが公開されたら、調整されていきます。アイデアを思いついたら、試して、公開します。ある論文から新しい知識を生み出しても問題ありません。標準的な学術プラクティスに従って、オリジナルの論文を引用しなければなりません、これが問題だとは思いません。これはスクーピングではありません。スクーピングというのは、他を押さえ付けて同じことを自分のものとして主張することです。プレプリントと同じく、既に公開したものに対して先取権を得ることができます。もちろん結果を再現して、問題なく再現できることを示すことのメリットもありますが、それが既に行われたことであれば、そのことをはっきりさせておかなければなりません。

●林 武田先生宛てにテクニカルな質問が来ているので、質問と共にお答えいただけますか。

●武田 はい。例えば ID についてですが、arXiv には arXiv ID というものがあって、arXiv コミュニティでそれを書誌に書くのです。ただ、arXiv ID は arXiv の世界では通用するし、つながるけれど、外の人には知らないわけです。だから、DOI がないのはデメリットです。

arXiv で DOI を付けられない理由は、システムとお金の問題があります。DOI は CrossRef だと 1 件 1 ドルなので、それ掛ける本数分のお金がかかるのです。また、対応するメタデータをきちんと提供しなければいけません。さらに、システム改修も必要なので、ま

だ DOI を付けられていません。

改修できないのはお金のせいかという質問がありました。一つはお金の面が大きいと思います。今、システム改修のためのお金を積み立てているのですが、それでも満足な改修はできません。30 年の歴史でシステムが複雑になっているので、システムを入れ替えて機能を変えるというのがなかなか難しいというのが二つ目の理由です。

arXiv の著作権については、arXiv は何も主張しないので、そのまま著者が持っているはずで。

メタデータは何が悪いかというのは、私もよく分からなくて、システム担当の人が盛んにそこは難しいと言っていました。

次は重要な点です。arXiv に優先的に出すのはなぜかという質問です。伝統的な物理学の arXiv は、まず速報を出して、いずれきちんと書いたものをジャーナルに出すという、ファーストステップとしての役割が大きいです。コンピュータサイエンスの場合は、とにかく早く自分がやったことを公開したいという目的があります。もちろんその後にジャーナルに出すこともありますが、ジャーナルに出す手前で出すというよりは、とにかくまず公開します。そこで重要なのは、自分のウェブサイトに出すよりは、arXiv というブランドがあるところに出す方が、みんなにすぐに気付いてもらえるということです。

昨日、日本のとあるコンピュータサイエンスの研究室に行って、学生向けにオープンサイエンスの講義をし、その中で、arXiv の話もしました。担当の先生が学生に向かって「みんな何かやるときに、みんなも arXiv を取りあえず調べるよね」とコメントしていました。コンピュータサイエンスでも arXiv を取りあえずチェックして、似たような研究がないかを調べというのが習慣になりつつあります。そういう意味で、arXiv に出すのには、早く、見つけやすい場所に置くというインセンティブが働いているのだと思います。

●林 坊農先生に「プレプリントサーバーに出して評

価されるのですか」という直球な質問が来ているのですが、周りの反応はいかがですか。

●坊農 評価されるということが何を指すのかにもよりますが、自分がやった研究がプレプリントサーバーを見ればあるというのは、自分は非常に気持ちがいいです。プレプリントサーバーに上げているということは、学術的には評価されると思います。けれども、それ以外の点で評価されるかどうかは謎です。私としてはアップしてあるからいいというところがあります。それだけです。

●林 評価には軸があります。これは悩ましい話ですが、博論審査などの場合は、査読付き論文を何本出さないと要件を満たさないというルールが存在しています。つまり、ドクターを取る前の若い研究者にはプレプリントは使いづらいのではないかと趣旨のご質問が来ているのですが、Rebecca さん、何かコメントを頂くことは可能ですか。

●Lawrence 若い研究者は大規模な研究を行う時間があまりありません。幸運にも重要な結果を得られる人もいるでしょうけれど、それ以外の人は違います。若手研究者が抱える課題は、小さな結果がたくさん得られても出版することがとても難しく、それでもキャリアを積み重ねるには過去 2~3 年間の成果を示さなければならないということです。ここでも、ペースが遅いジャーナルのプロセスがネックになっています。プレプリントや私たちのようなモデルなら論文を出すことができ、これまでに何をしてきたか示すことができます。もしうまく時間内に査読もしてもらえれば、専門家のコメントが得られるので、若手研究者がキャリアを重ねる上で本当に役に立ちます。

●林 今の点について何かありますか。

●武田 建前を言うならば、博士審査でジャーナルな

どを要求するのは、本来ならば審査委員会の仕事の放棄です。審査委員会が本来それを審査しないといけないうのに、外部の評価をもって代えるというのは、本当は仕事をしていないということになるわけです。

だから、本来であれば、論文の中身で評価すればいいということになるわけです。本来は補助的情報としてジャーナルを使うべきです。そういう意味では、もちろんそのようなジャーナルをポジティブには使えません。では、そうでない場合はどうするかというのは、それぞれがこれから考えていくことかと思います。

ちなみに、コンピュータサイエンス系は、ジャーナルに出すことには全体としてあまり価値を認めていません。私が所属しているところでは、コンピュータサイエンスのカンファレンスランキングで上位にある会議に参加すればよいとされています。こと博士論文取得に関しては、本来問題に戻す、ある種いいチャンスかもしれないと思っています。

●栗山 そうはいつても、われわれの大学院の指導の中でも、早くジャーナルに出せと言っています。文系でも社会科学だとそのような傾向があります。

あとは補足で、大変失礼なのですが、坊農先生はグリーン OA を少し誤解されているような気がするのです。グリーン OA というのは『Nature』や『Science』など査読付き雑誌に出したものをインターネットで公開することです。ですから、評価を査読付き雑誌に頼っているのです。従来型の評価システムが変わらなければ、オープンアクセスの一つの手段としてグリーン OA がもっと評価されてもいい気がします。

●林 学会がどのようにプレプリントの質を考えるかというのは非常に重要な話なのですが、日本はまだこの問題への気付きが足りないような状況です。

Rebecca さんに伺いたいのですが、学会 (learned society、research community) に所属する人たちとコミュニケーションは取られていますか。あるいはどのような内容の議論をされているのでしょうか。その人た

ちはその分野でプレプリントを含むクオリティコントロールを担っていると思うのですが。

●Lawrence 私たちは資金提供者だけでなく、研究者ともたくさん話をしていて、特に若手研究者はシステムの変化を渴望しています。彼らは目の前にあるものを好ましく思っていない。ボストンで始まった Future of Research などのイニシアチブがたくさんあって、幅広く拡散しつつあります。これはシステムを変えたい、新しいことを試したい、進めていきたいと考えるポストクのグループで、どんどん成長しています。私たちは明らかにシステム全体を変える必要があります。若手研究者だけでなく、もっとシニアの研究者からもたくさん学ぶ必要があります。問題は、評価を行うのは、現行のシステムで現在の立場を手に入れたシニアの研究者だということです。ですから、彼らにはシステムを変えるインセンティブがありません。プレプリントを認める資金提供者などはどんどん増えています。ウェルカム・トラストや NIH、他にも多くの組織が、プレプリントへの投稿を明確に推奨しています。でも、これらの成果に対する査読がないと、成果を目にすることはできても、その質がどうかは必ずしも分かりません。ですが、この状態は変わりつつありますし、次の世代は変化を熱望しています。

●林 学会、あるいは研究者とのコミュニケーションを今後変えていくためのという意味合いで、パネリストの方、何かコメントはありますか。

●武田 F1000 の活動で気になったのは、コミュニティとしての学会、コミュニティとしてのジャーナルというものが要らないということなのかということです。ライフサイエンスはよく分かりませんが、他の分野では、学会あるいは学会が発行するジャーナルが一つのコミュニティであって、そこに集まることがある種の質の保証につながると考えているのではないかと思います。

ます。そういうことはなく、もっとフラットで構わないという考え方なのでしょうか。それとも、F1000 も、あのプラットフォームの中でコミュニティをさらにつくってやっていくのでしょうか。

●Lawrence F1000 というプラットフォームはさまざまな領域をカバーしますが、私はコミュニティをなくすべきだとは言っていない。コミュニティは学会やソサエティと同じく重要なものです。私が提案しているのは、これらのグループに対するこれまでとは少し違う役割で、大事なことは結果の共有を他のことから切り離すことです。このような学術コミュニティや査読者、ソサエティはこれからもキュレーションをしたり、影響や質、重要性、新規性の評価を続けたりできます。私たちはただ、出版を可能にするためのジャーナルは必要ではないと述べているのです。arXiv は素晴らしい基盤で、arXiv をあらゆる分野に拡大し、そこに査読を加え、それからまた別にキュレーションを加えることができれば、完璧なものになり、その中でさまざまな分野のさまざまなコミュニティを持つことができます。

●林 学会関係はよろしいでしょうか。

●坊農 林さんの調査にあったとおり、生命科学は学会誌を無視していて、あまり相手にしていないので、学会でコミュニティの形成をしているという気はあまりなくて、どちらかというと同窓会に出ているだけという感じです。

●武田 それは元々、学会とジャーナルは切り離された形になっていたということですか。

●坊農 分野によると思います。私はどちらかというと、巨大な人数を抱えている生命科学系の学会にしか所属していないので、という意味です。

●武田 なるほど。生命科学だと学会の役割はそういう役割であって、その学会はジャーナルは自分たちで出さないのですか。

●坊農 大きな学会はもちろん持っています。そういうところに出す人もいますが、いろいろな複数の学会に所属していて、より規模が小さい、専門分野に沿ったところに出す人もいます。そういう場合でも『Nature』や『Science』に出してからという流れがあります。それは一昨年、話しました。

●武田 それは分野によってかなり違うのでしょうね。学会が出すジャーナルが重要な分野では、今 Rebeccaさんが話したように、学会がそういうことを主導して変えることもできるかもしれません。arXiv のときも、今回 ACM (Association for Computing Machinery) の人が Member Advisory Board (MAB) に参加していました。arXiv は、今までは基本的にメンバーは大学だったのに、今年から学会をメンバーに入れるという決定をしました。今まではどちらかというと大学が維持するというスタンスだったのが、学会もその維持のメンバーに加わる、これは大きな変化です。

実際、ACM と実験をやって、学会が連動する。特に学会は、「われわれは商業出版社と違って arXiv と協力したい、学会にとってはそれが良いことだと思っている」と言っていたのです。Rebeccaさんがおっしゃったような道は、学会がそのように思えばできるというのは、確かにそのとおりだと思います。

●林 面白いですね。プレプリントを軸に研究者の所作が変わり、それが変われば研究者コミュニティの所作も変わる。それはドメイン別で、全く横並びではないことだけは今、明確に分かりました。皆さんご自身のドメインに落ちるとどう議論が展開するのかというのを考えていただくようなフェーズになってきました。

では、今、学会の話になったので、そろそろ図書館の話にしたいと思います。プレプリントサーバー時代

が来たときに、図書館あるいは図書館員はどのように処すべきかという質問も来ています。この問いは SPARC Japan ができたルーツにも深く関連する論点です。図書館の方々の今後の活動に役立つ話について、栗山先生にまずコメントいただいて、その後、Rebecca さんにフォローいただければと思っています。

●栗山 図書館がプレプリントサーバーを維持するという点について、コーネル大学の図書館が arXiv を維持・管理してきたというのは非常に特殊なケースだと思います。普通の図書館は機関リポジトリということで、グリーン OA の受け皿になっていたわけです。今の出版の傾向が変わらない限り、グリーン OA というのは、あまりうまくいっていませんが、少なくとも存在意義はあるだろうと考えます。

もう一つは、図書館自体が出版事業に乗り出すということです。大学の紀要というのがありますが、日本の特殊なものだと聞いています。あれは大学がお金を払って、大学の教員の研究成果を発表する雑誌なので、言ってみれば、一種のゴールド OA です。そこできちんとした論文が出れば、一つ図書館の事業として成り立つということは言えると思います。現実問題として、たまにとんでもない論文が掲載されているので、そのあたりは責任を持たなければいけません。

私自身が経験したケースとして、ある論文が著作権を侵害しているのではないかという疑いがあったて、法律の先生に相談してみたら、やはりこれはまずいということになりました。そういうことで、直接の責任は紀要の編集委員会にあります。図書館がリポジトリとしてその論文を発表するということになれば、当然責任が問われます。そのあたりは腹をくくって、出版事業をやるのだというぐらいのつもりでやらないといけなくなるのではないかと思います。

●武田 これは全ての大学図書館ではないのですが、例えば、今でも日本の大学の一部で、サブジェクトリポジトリに相当するものを運営していたりします。つ

まり、自分の大学の教員ではなくても、例えば数学など、ある分野のものは、国内の誰でもここに置いていいですよという道も一つあるのではないのでしょうか。

●栗山 はい、もちろんそういうことも可能だと思います。その場合は、当該分野の学会と緊密な連携が必要になると思います。

●Lawrence 違うポイントですが、図書館が果たせるもう一つの役割は、研究者がさまざまなオプションについてあまり知らない、何ができて何ができないのか分からない、データ共有といった分野で、研究者を支援することです。研究者を支援するというこの重要な役割には図書館員がまさに適任です。また、もっと幅広い観点からは、このようなさまざまな問題をめぐるオープンサイエンス教育も重要です。どのようにすればいいのか、どんな落とし穴があるのか、どうすれば落とし穴を防げるのか、そして一般的な教育も含めて、図書館員にできる重要な役割が幾つかあります。

●林 ありがとうございます。フロアの方から何かコメントはありますか。

●フロア 1 千葉大学附属図書館の図書館員です。今日、一番衝撃的だったのは、どの発表を見ても「図書館」という字が全く出てこなかったことです。林さんのまとめスライドにも全く出てこなかったのは衝撃的でした。Rebecca さんのご紹介のようなプラットフォームが出てくると、図書館がこれからこの中でどういう役割をしていったらいいのか、大変疑問だったので、今データシェアリングのお手伝い、エデュケーションの部分で何かができるということを示していただいたのは、とても良かったと思います。

もう一つ、ある程度、質保証されたものを例えば機関リポジトリにアーカイブしていくという考え方があり得るのかを Rebecca さんにお聞きしてみたいです。

また、坊農先生にお聞きしたいのは、プレプリント

と機関リポジトリの違いのようなことです。機関リポジトリは、グリーン OA の論文だけを入れるわけではなく、プレプリントサーバー的な働きができるのかどうか、また、機関リポジトリに出したいと思うのか、なぜ bioRxiv の方に出したいと思うのかをお聞きしたいと思います。

●Lawrence 私が考えている方法は、論文とそれに伴うデータの出版は異なるタスクであり、全て一緒にまとめる必要はないというものです。やらなければいけないさまざまな種類のサービスがあります。品質保証のチェックもその一つで、これは図書館で簡単にできます。論文がプレプリントなどに投稿される前にすることだってできるでしょう。ありとあらゆるグループがこれを行うことができるかもしれませんが、図書館はその中でも適任でしょう。

同じく、データ共有も独立した一連のチェックとして行うことができます。今のところ、私たちの F1000 の編集チームはほとんどの時間をチェックと主にデータをめぐる研究者のサポートに費やしています。ここでやらなければいけない仕事は大量です。この仕事は全て、いろいろなグループにできることです。このサービスを提供するのは出版者である必要はありません。ポストドクのグループでもいいでしょう。より良いサービスを提供できれば、より多くの人がそのグループを使うようになります。

●坊農 それをやっていただけというなら、とてもいい試みだと思うのですが、実際問題、難しいのではないかと思います。bioRxiv に出して、2 日で出ましたと言いましたが、それぐらいのスピードでやっていただけるのかということがあります。また、bioRxiv はただ単にアーカイブしているだけではなくて、それを分野ごとに Twitter で流したりしています。そういうことが必要です。全世界に向けて、それだけのフォロワーが稼げますかということですね。bioRxiv はすごい数の人がフォローしています。そういうところが

苦しいのかなと思います。

そして、そこに行けばあるという検索性が非常に問題になります。今までも、bioRxiv はいいいのですが、PubMed で検索できないということで避けていた人も割といるようなのです。皆さんが使う検索で引っ掛けられる必要があります。データの大きさが大きくなるので、その点では各所で持っていただける方が理想的ですが、手間を考えると大変ではないかと個人的には思います。

●林 図書館に関係することで、何かコメントがある方はいらっしゃいますか。

●武田 私のスライドでも示したように、今いろいろな手段で論文を公開する手段ができています。もちろん分野にもよるのですが、分野によらない共通の手段もあります。ですから、何か論文を出したいときにどういう手段があるかを、図書館で教育してくれると、教員や学生はありがたいと思います。これは Rebecca のコメントと同じですが、特に若い人は他のチャンネルがあるということを知らないことがあり、それを図書館を通じて知ることができたら、ありがたいのではないのでしょうか。実際は、若い人だけでなく、普通の研究者でもそういうところを知らなかったりするので、自分が特化した分野では、出すジャーナルにどのようなものがあるか全部知っているのですが、少し外れると知らないことが多いので、その周辺を教えてくださいるのはとてもありがたいのではないかと思います。

●林 そうですね。過去の SPARC Japan セミナーの議論でも article processing charge (APC) の問題が出たときに、出すべきオープンアクセスジャーナルを慎重に選ぶために図書館がサポートできることがありましたねという話になりました。最近ではハゲタカジャーナルに出さないようにするための教育をしましょうという話が盛り上がってきています。今回の議論でさらに、プレプリントサーバーはどういうところがありま

すかと聞かれれば、レファレンスサービスに近い形で、研究者の教育としてライブラリアンの果たせる役割が大きいという感じになってまいりました。

今日は北海道大学で数学研究をしながら機関リポジトリと長く関わってこられた先生が会場にいらっやっています。そのようなお立場から、今までの議論について何かコメントいただくことは可能でしょうか。初期の機関リポジトリ活動において研究者として非常に深くコミットされた方です。

●フロア 2 北海道大学数学教室の研究者です。武田先生のご講演にありました arXiv のプレプリントを論文で引用するという文化は、数学・理論物理では普通に行われています。そのような背景の下で、引用文献に記された arXiv のプレプリントに、出版後に、例えば追記する形で、ジャーナルアートの DOI を付与する、そのような技術は、今の技術で可能なのかどうかを伺いたいです。

プレプリントという文化についてコメントいたしますと、例えば arXiv 以前、紙媒体で一つのプレプリントを数百部印刷して、それを世界中の研究機関に送って先願権を確保していたという背景がありました。そうすると、あるプレプリントを作成した研究者の周辺だけで、ある研究が発展する、そのような非常に限られたコミュニティが成立すると。そういう背景を解決するために、デジタル技術がうまく利用されている、そういう解釈もできると思います。そのような背景をできれば皆さんが理解された上で、プレプリントサーバー等の議論を共有したいと希望します。よろしいでしょうか。

●武田 arXiv は、一つの論文を出したときに、バージョンを付けて、バージョン 1、バージョン 2 と更新することができるので、バージョン 2 になったときに、例えばアクセプトされたものの DOI を付けることはマニュアルでは可能です。

でも、システムティックにやる方法が用意されてい

ないのです。それが bioRxiv だと、bioRxiv に投稿して、そこからジャーナルに投稿するというパスを踏めば、自動的にリンクが生成されるので保証されます。それはジャーナルと bioRxiv の両方にとってメリットがあって、きちんとリンクが確保されています。

arXiv はそのような自動的なメカニズムを持っていません。また、今、メタデータに関してもそのようなものを持っていません。それで、OA のエコシステムの中に入っていないという批判を受けています。本当は直せばいいのだけれど、arXiv は 30 年の伝統があって、なかなか大変なようです。今、実はページを表示しているシステムは入れ替わったのですが、見た目は全く変わらないのです。それは多分、伝統を重視して、見た目をほとんど変えないようにデザインしたと。そこは arXiv の歴史から来るある種の足かせにはなっているのではないかと感じました。私はむしろ arXiv のユーザとしては新参者なので、そのように思いました。

●林 そういう意味だと、進化論的にはプレプリントサーバーも細分化していくことになるのでしょうか。バイオ系の新興のものだと、そういうしがらみにとらわれずに新しいデザイン、あるいは UI/UX を検討できる余地があるように思うのですが。そんな兆しはありますか。坊農さんは arXiv と比較されたことはないですね。そのあたりでお気付きの点はありますか。

●坊農 生命科学系だと、もちろん bioRxiv だけではなくて、PeerJ に出すとか、BioRN に出すとか、そういう選択肢はあると思いますが、bioRxiv に出しているのは、arXiv の伝統を引き継いでいて、ここだと皆さんが見てくれるだろうという願いがあったのだと思います。そこでまたいろいろ出てくるとややこしくなりますよね。だから、あまりそこは考えたくなくて、そこからジャーナルに出すということも伝統的にあるから、そこから違うジャーナルに行くという意味で、プレプリントの入り口としてはそんなにたくさん出てこ

ないのではないかと、私は思っています。

●林 そうですね、逆に今の議論から想起されるのは、プレプリントサーバーのデファクト化を実現する上で誰がその訴求力を持つか。始める人がすごく重要になってくるのでしょうか。そのコミュニティの中で名が売れた人が、必ずしもそのドメインのハイインパクトな研究者でなくても、何かしらの求心力を持った方が始めれば必然的に集まってくるように思えるのですが。

●坊農 そうですね、私は生命科学でもだいぶマイナーな分野でコンピュータを使う実験をしています。生命科学は実験して自分のデータを発表するのが論文誌なのですが、そういうデータを持っていない中で論文を書いてきていて、そういう意味では、なかなか論文を受け入れてくれない分野なのです。そういう人間が出すに当たっては、コンピュータサイエンスの arXiv に出すのと同じような感じで、bioRxiv を利用したいと思っています。

今日お話ししたものに関しては、査読誌に出すことは考えていなくて、あそこで止めておこうと個人的には考えています。そこから出てきたデータを使って、何か新しい発見をしていただくというように持っていきたいと思っています。私は bioRxiv 本来の使い方ではない新しい使い方をしている、そのようなやり方が実現できる場として大事だと思っています。

●Lawrence もっと多くの人にプレプリントを使ってもらよう促すという点では、資金提供者がある程度支援してくれるだろうと思いますが、ジャーナルに掲載される標準的なナラティブの研究論文を越える別のタイプの成果が資金提供者の資金で行われたものとして認められ、次の資金を獲得したり次のキャリアに進んだりするときに認められるようになったら、大きな差が生まれると思います。これは、NIH などがプレプリントを認めると言い出したときにある程度見られました。突如として、よく知られた研究者の小規模

なグループがプレプリントしていたときよりも、はるかに急激な成長が起こったのです。最終的には、所属機関や資金提供者から、プレプリントだけでなくデータや方法、資料のシェア、それから査読活動や授業といった、学術エコシステムにとって非常に重要な幅広いあらゆる活動に対する報奨がなされ、インセンティブが与えられるようになります。

●林 ありがとうございます。今回は、議論を明確化するために研究論文のプレプリントにフォーカスして企画したのですが、いみじくも Rebecca さんの最後のコメントで、プレプリントを考えていくということは、論文以外にも研究成果の対象を広げてサイエンスをもっと発展させていきたいと思いますという趣旨であり、すなわちオープンサイエンスを進めましょうという話になりました、非常にありがたい最後の締め言葉を頂いたと勝手に解釈しております。本来ならパネリストの方に一言お願いしたいところではありますが、ちょうどお時間になりました。これにてパネルディスカッションを終了したいと思います。改めまして、パネリストの皆さまに拍手をお願いいたします。どうもありがとうございました。