

講演会「オープンアクセスの今と未来」

日時：平成 31 年 2 月 18 日（月）

質疑応答・意見交換 15：50～16：30

コーディネーター

逸村裕氏（筑波大学図書館情報メディア系教授）

登壇者

山地一禎氏（国立情報学研究所 オープンサイエンス基盤研究センターセンター長）

林正治氏（国立情報学研究所オープンサイエンス基盤研究センター特任助教）

尾城孝一氏（国立情報学研究所オープンサイエンス基盤研究センター特任研究員）

逸村氏：質疑応答の前に、壇上の国立情報学研究所の山地先生、自己紹介とオープンサイエンス基盤研究センターの活動についてお話しください。

山地氏：NII の山地と申します。センターでオープンサイエンスの基盤を作っています。

逸村先生のお話しの中で、どのようにオープンアクセスを実現するか？という話がありましたが、全世界の人がエルゼビアやシュプリングーなどの海外出版社を国レベルでボイコットする動きがあるのは、一つは彼らの利益率がすごく高いんですね。大手三社ともに 30-40 パーセントある。この高い収益率をもっているところが、学術者として考えていかねばならないところです。いずれにしろオープンアクセス、論文を書くのはお金がかかる。誰かがやらなければならない。大手出版社が悪という訳ではない。うまくバランスをとるのが一番大きな課題だと思います。その時に論文はゲームチェンジするのが難しい状況にあるんですが、研究データは国のアセットとしてきちんと日本国内に組織として研究者が研究の中で作ってきたアセットを保持して管理していくという機関を作っていきたい。この中で我々は各大学の共通基盤として、各大学が競争領域として個々にやるのではなく、システム開発にしても、運用にし

てもお金がかかるので、協調領域として共通基盤を作るということで、我々センターが担当してやっています。

従来は機関リポジトリや、CiNii などの検索システムが主だったのですが、現在、我々は研究基盤の管理システムというところで大きくジャンプしようとしている。簡単にいうと研究データ管理のためのドロップサービス。これは難しいサービスで 24 時間 365 日、決して止めてはならない。チャレンジングだがうまく研究者の方々に認めていただけるものにできれば一緒にオープンサイエンス、オープンアクセスを考える切っ掛けになると考えています。

「大学のオープンサイエンスは図書館が主導すべきとお考えですか？」

尾城氏：オープンサイエンス自体はサイエンスなので図書館員が直接オープンサイエンスを行うことはないと思います。但し、オープンサイエンスを支える研究データ管理は図書館が主導すべきだと考えています。しかし、図書館単独ではできないので学内の関連する部署と連携しながら、図書館が主導するべきだと考えています。

「研究者のデータ作成時にデータライブラリアン

が関与しないのか？」

尾城氏：実際にデータを生み出すところに図書館員は関与できないと思っています。しかし、研究プロジェクトで生み出された生のデータをそのまま共有したり、公開する訳にはいかないのです。加工する必要がある。それはデータのキュレーションと呼ばれることがあり、それには図書館員も関与できると考えています。但し、その時に大きな問題があって、キュレーションをやろうとすると研究分野の知識が必要になってくる。

アメリカの大学図書館には、伝統的にサブジェクトライブラリアンという専門の分野で学位を持っている人がいる。今アメリカの図書館ではそういう人たちがデータキュレーションの仕事の方にシフトしてきている。日本ではサブジェクトライブラリアンという伝統はなかったので、図書館員が単独でデータキュレーションの仕事を担うのは難しいかもしれません。図書館だけでなく研究者の力を借りながらキュレーションに手を出していくべきだと思っています。

これからのことを考えると、最近、大学図書館に採用された人たちの多くは専門分野の知識を持っています。いろんな学部とか大学院を出て、マスターやドクターの資格を持っている人たちも来る。図書館の若手職員というのは分野における専門知識を持っています。それを図書館の現場で、データキュレーションという仕事の中で活用できるよう考えるのがいいのではないかと。今後、専門知識を活かしたデータキュレーションができる図書館員が生まれてくるのを期待しています。

山地氏：一番初めの質問「図書館が担うのですか？」は、図書館以外の人がしているとは思えないのですが、図書館の人がこの質問をしているのだったら、主導してほしい。我々は今JPCOARでトレーニングコースを作っているんですが、研究管理を今一番勉強してわかっているのは図書館。我々がサービスを作っていく上

でも研究者がどう使っていくか？というユーザーストーリーを常にベースにシーズを押し上げていく形で作っている。図書館がわかっているなら、面白いところがあるならどんどん出てきてくればいいし、筑波大学の図書館は管理職も含めて若手をサポートしてくれる状況にあると思うので、ぜひ活躍の場としてほしい。

オープンサイエンス基盤研究センターではトレーニングコースを作っていて、システム、コンテンツを作り、サービス提供もしている。図書館の方々に国立情報学研究所に来て、大学に戻って使えるようなコンテンツ作りを一緒にするのも歓迎したい。

逸村氏：補足ですが、自己紹介で文部科学省の科学技術・学術審議会 学術情報委員会の話をしました。この2年間、日本のいろいろな分野の研究者にヒアリングをしています。

委員会の議事録は公開されていますし、資料も相当部分、公開されているので見ていただけるとよいのですが、既に日本の学術の世界においてもデータの公開を行っているところが分野ごとにかかなりあります。理系も文系も含めてです。

例えば先だってあったのが政治学。選挙のデータに関して、大量のデータを公開している事例がありました。国際的に協力しながら公開している例もあります。その中で日本が先頭を走っているのが地震のデータですね。東日本大震災の時のデータをとっており、公開しています。そういうものに関しての管理についても表現されているものがある。

事例をみてその中で図書館員がどこまで対応できるか。図書館員が対応できない巨大なデータや主題に基づいたメタデータについてはそこで図書館員は何ができるか。そこらへんに関わっている研究者が自分の大学や機関にいるということについてデータをとる必要があると思います。

これに関しては、尾城先生が先ほど紹介して

いましたが、イギリスのエジンバラ大学データキュレーションセンターが先進的なことをやっています。図書館に付随しているところで、その担当者と話した時に巨大なデータは学会や研究所が確保している。大学が何を持つかというのはポリシーの問題であろうということで、なんでそんなに頑張るのと聞いたら、エジンバラ大学は今世界ランキング20位で、そのランキングを保つのにその作業は不可欠だと言われました。

山地氏：有名どころの大きな大学でe-サイエンスとかデータセントリックスサイエンスが2011-2012年頃に始まったところに研究データというものに組織として重きをおいてそこに投資していく。研究データ管理とかデータ解析、今でいうデータサイエンスですが、それについてきちんとサポートして組織を作って若手を育成してということが大学のブランディングにつながるということを経営層が把握した上で、大学の舵をきっている。エジンバラ大学がそのいい例で、その中で研究データ管理、データサイエンスを組織として捉える。大学の検証の中にもそれを捉えていくという動きができてきている。今、我々日本ではそこまでの大学はない。データセントリックスというところから行くよりも、オープンサイエンスとか、もっと飛び越して欧米の動きに追従していくなど、大学のブランディングを強化するために大学として何をするか考える非常にいいチャンスだと思います。

「リポジトリ登録に協力的になれない研究者は多いのでしょうか？また協力的になれない理由は何でしょうか？」

林氏：図書館員から聞いた範囲での回答になりますが、リポジトリが研究者から見えていないと思っています。オープンアクセスのためだということも最近は伝わり始めているようだが、これまではそれが見えていなかったし、研究者

が主体的に動くという形になっていなかった。

研究者による登録がうまくいっている機関の事例を見ると業績管理の仕組みと紐づいているなど自身の評価に関係するところがあると研究者は登録するのではないかと思います。現状、いろいろやることのある研究者にとって登録は優先順位的に落ちているのではないかと考えています。

研究データを登録するということができたときに、機関リポジトリの役割が大きく変わると考えています。オープンアクセスの受け皿、紀要論文など機関のオリジナルのコンテンツの発信元という役割から目的が変わってくる。

研究データとなってくると研究者を引き付ける何かが必要だと思っています。機関がリポジトリを何をもって公開していくかという位置づけが重要で、それによって研究者は動くと考えています。

山地氏：インセンティブは重要な問題で、研究者は金か自分の昇任か人事というところでしか動くインセンティブはない。オープンアクセス、オープンサイエンスはインスティテュートリサーチに紐づけて、人事評価に回ってくる仕組みとしないと元気が出てこない。

オープンアクセスの時はジャーナルをオープンにするだけだったので、大きな変化はなかったが、データとなると研究者が皆初めから終わりまで研究しないと評価されないのではなく、ポーションポーションでも研究者が評価されるような、データ取るだけ、解析するだけ、それらを失敗データを含めてどんどん公開する研究者が評価されるような仕組みにつなげられるような制度設計とかサービス、サービスが評価につながるエビデンスを出すと思うので、そこに注力しながらシステムの開発を並行しながらやっている。

「APCを支払うのが前提で改革が進むと、APC

を払える財力のある研究者、機関のみが研究成果を発表できるのでは？」

逸村氏：メガジャーナルにしても PLOS ONE が 1,350 ドルという話をしました。これが Nature Communications ですと 5,000 ドルを超えて、金持ちでなければ出せない。日本のオープンアクセスジャーナルへの採択状況を調べると一位が東大、二位が京大、三位が理化学研究所という研究費をたくさんもっている機関の順位になっている。この状況は学術情報委員会でも議論になっていて別の学術分科会でも議論になっています。研究そのものをオープンアクセスにする、オープンサイエンスにする、データを出す、ジャーナル論文を出すというところで初めから研究費の一部を間接経費として論文の投稿料やデータの搭載に充てるべきという議論が進んでいます。実現可能性という問題がありますが、APC が払える機関だけが独占している訳にはいかないことについては議論しています。

山地氏：逸村先生のところで SCOAP3 というプロジェクトの紹介がありましたが、高エネルギー物理の人たちは arXiv というプレプリントサーバに論文をプレプリント、出版社に出す前にそこに登録して、査読に時間がかかるので自分が行った研究だということを押さえる文化がある。そこから全世界のデータを取ってきて、著者がどこの国に属するか、全部、名寄せをするんですね。その後に各国に対して あなたのところは high energy physics にこれだけ論文が出たからこれだけ払いなさいという分担金を CERN が中心になって組織を作っているのが SCOAP3 です。図書館の活動の中だけで SCOAP3 にコミットして来たんですが払えないんですね。図書館が払っていた分だけでは払えない大学も当然出てきて、結局、研究者が論文を買う、論文を出すというのはお金がかかるというのを理解したうえで、図書館の今までのサブスクリプションだけでは払えない部分をどう担保するか、研

究者と一緒になって考えなければならない。その中で、逸村先生がおっしゃっていたように、研究費の中で、直接経費または間接経費の中で払うという話があったんですが、論文は研究が終わってからでるパターンが結構多いので、組織として吸収せざるを得ない。組織の中で年金モデルみたいにある程度プールして論文を出していくという仕組みを作っていく必要があって欧米はそういうケーススタディが何年も前から進んでいるんですけど、日本ではそこまでまだ辿り着けていない。という現状を何とかしなくてはならないというのがありますね。

逸村氏：この辺に関しては NII のサイトの SPARC Japan というところで資料を公開しています。

日本はそういう意味では払うべき拠出金を払いきれていない。今の話の流れでお分かりだと思いますが、このところ日本の科学論文の数が減っているじゃないか、という話をお聞き及びかと思えます。然るに high energy physics は減っていないんですね。それ故に払うのが増えてしまっ、ある部分 CERN に払ってもらっているという事態になっています。

「原則的なデータ管理の永続的な保存管理の責任はだれが負うべきか？」

尾城氏：大学が組織として責任を負うべきだと私は考えています。大学がこのことをポリシーとして謳っていく必要があると思います。

逸村氏：私の講演の最後に少し PLOS ONE の経営苦境の話をしました。機関リポジトリが潰れるのと、オープンアクセスメガジャーナルが潰れるのとどちらが早いかなという話はしたくないんですけども、これは信頼というかインテグリティ(尊厳)の問題として重要だと思っています。

大学がこの先どうなるか分かりませんが、どこかの大学がどこかに吸収されるとしても、そ

の際にリポジトリなどは資産として吸収するということは政策的に重要だと、これも議論はしていて理解はしているというところです。NIIのJAIRO Cloudの立ち位置というのは複雑なものを背負い込んでしまう訳ですが、山地先生その辺はいかがでしょうか。

山地氏：それはありますが、我々もサービスを提供する上で受益者負担モデルをやっていくとなると、大学として永続的というものをどういう風に捉えるかというガイドラインなりポリシーを決めるというのが難しいですよ。10年保存と言っていますが10年も保存できていないし10年でいいのかというのは大学として考えた時にその資産をいつまで残すのか、残すのはお金がかかるのでどの部分は忘れてもいいのかというのを研究者主導でガイドラインを決めていく必要があるし、分野ごとに違いますよね。それをどういうふうにするかは大きな課題だと思います。

「ヘルスサイエンス系の研究者の方から、PLOSに投稿するときにデータの公開を求められます。現状でデータをアップロードできるデータリポジトリはどこにあるのでしょうか。」

尾城氏：PLOSはデータを公開するための推奨データ公開リポジトリのリストを挙げているのですが、このリストは2種類に分かれていて分野別のデータリポジトリと、分野を超えた汎用的なデータリポジトリですが、日本のデータリポジトリはそこに含まれていない。分野別のデータリポジトリの中に一部日本の機関も協力しているリポジトリはあるんですが、運営の主体は全て海外の機関ですから国内のリポジトリは今のところゼロとっていいです。このままではデータの海外流出がどんどん進んでしまう訳ですね。ですから機関リポジトリにもデータリポジトリとしての機能をきちんと持たせていかないといけないと考えます。

林氏：ヘルスサイエンス系のところは存じ上げていませんが、国内のデータリポジトリの運営者達はそこに危機感を覚えていて、いかに国際的なレジストリに登録してもらうかを目指してどういった体制を整えていけばいいかデータをどのように保持していけばいいか検討、議論始めているというところです。

逸村氏：既にご承知のとおりSTM系はほとんど海外のジャーナルに論文を押さえられていて、日本の学会が出している英文のジャーナルも多くは海外の出版社経由で発行されています。日本のお金で日本人がやった研究も、日本のお金で外国人が日本の学会で出したものも全部海外から買い戻すような状況になっています。それが海外流出ですが、これに関しては深刻に考えています。

それに加えてデータもというのは、まずいという話はしていますし、何がまずいかというと今度はそれを評価するという視点を失ってしまう、全部海外に行ってしまうというのが議論のあるところです。

「リポジトリにメタデータを登録するときにどのような自動抽出を考えているか？」

林氏：PDFから自動的にメタデータを抽出して、メタデータを保管するというのは実装できていません。ソフトウェアによっては例えばPDFにもメタデータが埋め込まれているものを利用して登録するというのは可能だと思っています。テキストを解析して抽出するということもあるかなと思いますが現状ではできていません。

「日本語論文の自動英訳の開発状況は？」

林氏：自動英訳についての開発は進んではないんですが、日本の論文、紀要等を海外に見てもらう時に英訳は重要だということはお出しています。現状はできていませんが、今後の課題と思います。

山地氏：自動英訳の機能を我々がお金を取ってきて開発するかというと、しないんですね。そういうのをできるだけ周辺のサービスとして、疎結合をしながら、API を叩きながらくっつけてくれるシステムのアーキテクチャになっているので、例えば PDF からメタデータをとってきたりとか、実験装置と連携しながらメタデータを取ってくる、分野特化の機能となるかもしれませんが、共同研究と一緒にやりながら大学と一緒に作っていければと思いますのでいつでも声をおかけください。

「Plan S の現状及び今後について」

逸村氏：現状、日本語で書かれてわかりやすいものは、昨年、「大学図書館研究」9月号にこの手の話をまとめています。JUSTICE 事務局長が書いている記事もあります。ドイツのマックス・プランク研究所の Ralf SCHIMMER さんが書いたものを日本語に訳したのも載っています。そこら辺が基礎的資料としてお読みいただければと思います。

NII のもとでやっている SPARC Japan でもイベントとして年 4.5 回オープンアクセスがらみのことをやっています。その中で Plan S に関しての話、ほかの今一連の話になっておりますオープンサイエンスのデータ公開を含めた話も出ております。この辺の資料も公開されておりますし、ある部分は出演者の動画も載っています。こちらもごらんください。

簡単に短く読めるものでは国立国会図書館が編集しております「カレントアウェアネス」でも情報として雑誌記事、WEB サイトの紹介をしています。学術情報の専門家が編集しておりますので手短かに手に入れるには、こちらをご覧ください。ただけるといろいろな情報が載っています。また、JST の方でも同じようなサービスがあります。いずれにしても毎日のように新しいニュースが出ています。

「現状、大手出版社と海外の状況はどうなっているのか」

逸村氏：今、ドイツが巨大だということで、マックス・プランクが先頭を切っていますが、エルゼビア以外とはある段階までは妥結しているという流れになっています。しかし、これもころころ変わる。

ここら辺の論文のオープンアクセスがらみは 20 年以上長い闘いがあり、学術のオープン側が負けてきたんですが、最近はオープンサイエンスという文脈も含めて新たな戦争といってもいい状況で、その一翼を日本も担うという形で動いています。

実際ヨーロッパでは 12 か国の研究助成団体がスクラムを組んで戦うという流れになっています。アメリカ化学会、ACS ですね、などはジャーナルの論文などは条件をつけていますが、データに関しては論文に付随してデータをオープンにすることは OK を出しているという流れがあります。

徐々に動いている中で日本がどう立ち位置を維持するか、今、迫っている課題の一つです。

山地氏：Plan S ですが 2020 年 1 月がイミディエイトですね。テクニカルガイドラインもできていますが、厳しい条件で、現状、Plan S の提示するガイドラインにコンプライできるのは 15 パーセントくらいの雑誌しかない。STM で。人社系になると 3 パーセントくらいしか機能要求を満たせないという統計もあるので、1 月になってどうなるのか。かつプラットフォームが Plan S に対応しようとするとお金もかかる。お金を誰が払うか？というと結局アカデミックなところから流れていくしかない訳で、混沌とした状況がしばらく続く気がします。

逸村氏：補足しますと、ガイドラインが、先ほどお配りした資料に書いてありますが、今年の秋に公開されました。締め切りは過ぎています、出版社に対してフィードバックを求めています。

ます。出版社から順次回答が返ってきています。諸手を挙げて賛成しているのは PLOS くらいで他のところは条件をつけたフィードバックを寄せていて、それも全てオープンになっています。先ほどの私の資料に Plan S の関連 URL を書いておきましたので、ごらんください。

「機関リポジトリの将来についてご意見をお聞かせください」

尾城氏：これまで機関リポジトリには研究成果としての論文を載せ、論文の提出を研究者に促すのが図書館職員の仕事だった。今の段階では論文を集めるということに労力を注ぐよりも、研究データを収録することに労力を集約していく、それをうまく学術的に活用していくにはどうしたらいいかということ考えた方がよいのではないか。研究データに力点を移動することが必要、というご意見をいただいています。

論文の出版は研究ワークフローの最終段階に位置していますが、大手の商業出版社はワークフローの上流に遡って、支配を強めていこうとしている。上流でビジネスしようと考えている訳です。

最終的にジャーナルに載った論文ではなく、プレプリントやデータが彼らの主戦場なんですね。ですからエルゼビアは SSRN という社会科学系のプレプリントサーバを買収したり、Mendeley Data のようなデータシェアリングのためのプラットフォームを作って、それを売ろうとしています。

そういうことを考えると、個人的には機関リポジトリは研究ワークフローの上流にあるようなコンテンツを集めていくということを考えてもいいのかなと思っています。

林氏：先ほど山地からも話がありましたが、IR 的な評価に繋げることがないと研究者は動いてこない、もう 1 つはそこにあるデータを活用する何らかのサービスを作っていくことも必要で、両側面が必要だと思っています。サービス

を作るというのは機関が考えることができればいいのかもしれませんが、研究者のコミュニティから働きかけるとよいのではと思っております。

逸村氏：機関リポジトリは、2005 年から本格的にあちこちで作り始めました。

大きな成果として大学図書館がそれまで、外部で生産された図書や雑誌を、図書館を經由して利用者に提供するのが基本形だった訳ですが、リポジトリは学内の研究者の生産物を提供する、流れを逆に変えたところにあります。

リポジトリが普及しだした時に、あちこちの大学図書館にインタビューに行きました。その時にレファレンスライブラリアンからリポジトリは自分たちの業務に関係ないと思っていたが、リポジトリができたことによって、読む人が限られていた紀要が電子図書館機能によって、サーチエンジンによって学生にも容易に目に触れるようになってきた。その結果、授業を受けている、学部学生に至るまで、事前に先生を検索して、リポジトリに載っている先生の紀要論文を読んで、今回の授業の話はこれだと理解を深めていった。そこから寄せられるレファレンスサービスはより高度なものになっていったという言質を得たことがあります。このように学術情報のフロー、評価は変わるのであらうと思います。

IR、URA の活動も含めて学術情報の生産と流通という観点から、評価の問題と複雑に絡み合って発展させていくという流れになるのではないか。Institute Repository、あくまで大学のリポジトリです。なんで大学図書館が Institute Repository をやっているのかというのは、2005 年に NII で公募したんですが、1 大学を除いて、すべて図書館が応募してきた。その図書館以外の応募はよい提案ではなかったのでリジェクトしたという経緯もありました。

ここまで来て、NII でも JPCOAR をはじめとしてそれを支える体制を強化しています。ぜひ

図書館に留まらず、大学としてのさまざまな機能の一部である、これを高度化していかなければいけない、という流れでご理解をお願いいたします。

最後に、先生方お1ずつ人ずつお願いします。

山地氏：データか、論文かというよりも、オープンアクセスとか機関リポジトリというのが図書館の人たちのプレイランドというか、図書館の図書館による、っぽかったんですけど、今一度研究者が何を必要としているのかというところに寄り添って、そのニーズをきちんと汲んであげられるようなサービスや機能を考えるよいチャンスになるかと思いますので、論文じゃなくてデータになって、もうヘドロになってもいいと言っているんで、ぜひとも頭を柔らかくして一緒に新しいリポジトリシステムを作っていければと思っています。

尾城氏：図書館の人たちへの期待を述べさせてください。大学図書館の人たちはこれまで、柔軟に新しい時代の変化に対応してきました。図書館の電算化についてもきちんと対応してきた

し、電子ジャーナルの導入についても値段が上昇するという問題がありつつもアクセスできる電子ジャーナルの数は増えているので、うまく対応してきたのではないかと思います。図書館の人は瞬発力がちょっと弱いですが、持久力があるので、研究データとかオープンサイエンスについても、いったん動き出せば、持続的にきちんとした仕事ができるはずだ。オープンサイエンスやデータリポジトリについても相応の役割を果たすことができると思っています。

林氏：今リポジトリのソフトウェアを作っていますが、今後、データが入ってきた時にどういったサービスが提供できるか、研究者のニーズを捉えてやっていく必要があると思っています。ぜひとも皆さんのご協力をお願いします。それをサポートする仕組みづくりも一緒に頑張っていきましょう。

逸村氏：大学の機能が何かを問われている。図書館も URA も事務系も含めて、どのように大学が学術情報に向き合っていくかというふうな話が展開していけばと思っています。