

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？」

ディスカッション

堀井 洋	(一般社団法人 学術資源リポジトリ協議会)
林 和弘	(科学技術・学術政策研究所)
南山 泰之	(国立極地研究所)
山下 俊介	(京都大学 宇宙総合学研究ユニット)
武田 英明	(国立情報学研究所 情報学プリンシプル研究系)

●堀井 このパネルディスカッションは、「グリーンコンテンツの拡大に向けて」をテーマに、科学技術・学術政策研究所の林さんに加わっていただき、議論を進めたいと考えています。よろしくお願いいたします。

図1は、今回の SPARC Japan セミナーの企画ワーキンググループ (WG) の中で考えた全体のイメージ図です。左側が学術情報 (データ) の生成と検証・確定の部分です。論文や研究データなど、範囲をどんどん広げていって博物館資料まで広げていくことを目指しています。それがリポジトリや ID System JaLC などの横断的な情報管理体制モデルの構築につながり、それらを介してコンテンツを研究に利活用して、さらにその成功をフィードバックするというイメージを WG の中で共有しています。

山下先生と私は学術情報の生成に、南山さんと武田先生は横断的な情報管理体制モデルの構築に携わる立場です。南山さんには、コンテンツの利活用にもぜひ

とも入っていただき、国内外の動向や総括に関しては、林さんからいろいろご意見を頂きたいと考えています。

まず、林さんと私から、これまでの活動の紹介をしたいと考えています。初めに林さんからお願いします。

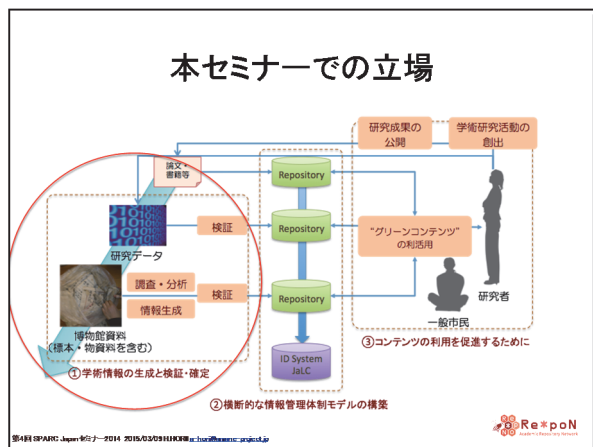
話題提供：オープンアクセスからオープンサイエンスへ 俯瞰と政策的要点

●林 私がこれまで行ってきた活動の中から、今までの話題提供者とは少し違い、政策的な観点、含意について簡単にご紹介します。

電子化実践、事業化から調査、政策研究へ

私自身は、オープンからほど遠い、科学の中でもオープン化に保守的な有機合成化学の研究に東京大学で携わっていました。その一方で、昔からコンピュータオタクだったものですから、研究室の Macintosh の管理をしているうちに、日本化学会の電子ジャーナル化に 1994 年ぐらいから携わることになりました (図 2)。日本化学会は世界的にみるとそれほど大きな学会ではないので、投稿して査読して組版して Web 公開するという、電子化の流れをひととおり、自らがイニシアチブを取ってやらなければいけませんでした。

最初は投稿査読システムの開発から始まりました。今でしたらコンプライアンスやセキュリティで問題になりそうなシステムだったかもしれませんが、1990 年代、2000 年代はまだそれが通用していた言わば牧



(図 1)

歌的な時期でした。それをやり終えたら、今度はビジネスにしなければいけないということで、事業化に取り組みました。そして事業化したら宣伝をしなければいけないとなり、チラシを作って国際会議でブースをつくって配りまくっている、そうこうしている間に、結果的に、電子ジャーナル事業全体を俯瞰することができるようになりました。俯瞰すると、まず、日本化学会の電子ジャーナルがどうしたら良くなるかという興味が出てきたのですが、当然その先に電子ジャーナルの未来はどうか、電子ジャーナルの先の学術コミュニケーションとは何だろうかという話になります。そのような流れで、学会や大学図書館、出版社の将来はどうかという文脈で興味が広がり、こうして SPARC Japan に私が参画させていただいていると理解しています。

そして現在、科学技術・学術政策研究所 (NISTEP) に 2012 年から勤めております。前職の日本化学会では電子ジャーナルの開発をし、併せて都合 6 年間 NISTEP の客員研究官を務めて、現在プロパーで働いています。NISTEP では研究評価や政策、大学、URA の在り方などにまで興味の範囲が及びます。学術情報流通の電子化実践と事業化をしつつ、調査、政策研究へ結び付けようとしています。科学の研究者はどういう人であるか、IT を使うとはどういうことか、コンピューショナルシンキングとはどういうことかということを、自分の拠りどころとして調査・研究活動を行っています。

(図 2)

歴史に習えば

今日のお話は、学術情報流通の歴史から見ると、昔は手紙でやり取りしていた研究者のコミュニケーションが、グーテンベルクのおかげで写植機が生まれて大量印刷になり、また物流ロジスティックスで公開の最適化が働いた。この時代がしばらく続いた後 Web が到来しました。どうやらその先には Web ベースの情報流通コミュニケーションの活動があるというのが見えているけれども、我々には PDF ファイルというものに象徴されるように、紙を電子化しているけれども、完全に Web ネイティブではないという情報流通の時代です。この状況に良くも悪くも甘んじている、今の人たちが満足していれば当面大丈夫という、科学者で言うところのトランジションステート (遷移状態) にいると理解しています (図 3)。

それをいかに右側にある将来の安定なところまで押すことができるかという人たちが、ここに集まってきているのではないかと思います。最終的には、グーテンベルクの時代と物流、要するに大量印刷と物流で情報の拡散が最適化されるデザインから、Web ベースのデザインに移っていく、この大きな流れがまずあるのではないかと考えています。

Open Access to Open Date and Open Science

図 4 は、今、内閣府でお手伝いさせていただくときによく使うマップです。2000 年代と 2010 年代の大きな流れの中、仮にオープンアクセスからオープンサ

(図 3)

イエンスへの流れというベクトルを書いてみます。図書館の人たちが集まると、まず話題になるのはオープンアクセスで、「アーティクル」へのアクセスです。もちろんブダペスト・オープンアクセス宣言では、再利用等々、もっと書いてありましたが、現実には、電子ジャーナルへのアクセスをどう活用するかという話でした。この流れから、オープンサイエンスになることによって「研究成果」をよりオープンにするという話になります。研究成果を再利用可能な形にするという流れでオープンサイエンスが広がりつつあり、さらに言うと、「研究活動」そのものを、よりオープンにしようという流れになっていると思います。

図書館活動としてセルフアーカイブや機関リポジトリの取り組みや、スティーブン・ハーナッドを筆頭とする研究者の活動がある一方で、出版社の方は、オープンアクセスゴールドジャーナルやメガジャーナルというものを作ってオープンサイエンスに対応してきているように位置付けることができます。その裏にサイエンスコモンズや、サイエンス 2.0 のような考え方があり、研究のパラダイム自体が変わっていくということを考えていくと、当然データも視野に入ってくるようになります。

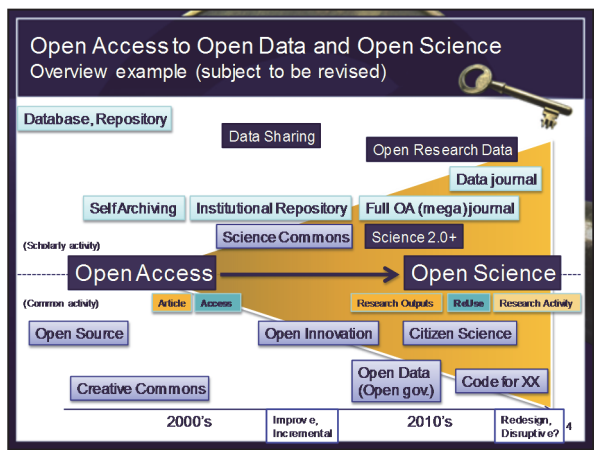
一方、研究データ自体は昔からあるわけで、ある先生に言わせれば原爆を開発していたころから研究データの共有はされていたという話もあるのですが、いずれにせよインターネットの整備等にかかわらず、データの共有は古くから議論されてきました。それをシェ

アしようという話は、インターネットの前のデータベース、デジタル化のところで起きました。そこからだんだんオープンリサーチデータのような考え方が生まれてきて、ここも出版社の方が、「では、データジャーナルを作りましょうか」ということで対応しています。

さて、ここまではどちらかというと科学者、研究者のコミュニティの中の流れなのですが、実はオープンサイエンスの持つポテンシャルは、公共的な文脈が含まれているということにお気づきの方もいらっしゃると思います。例えば、オープンソースの考え方があり、広く公共的に、ギークがいろいろな面白いコードを書いて「これはどう？」と皆で批評していたのです。

そこと並行する形で、芸術などを含むクリエイティブな活動を促進するために情報をどんどん共有しようという考え方からクリエイティブコモンズも生まれていました。そこからだんだんオープンイノベーションという考え方、よりオープンな情報、作品、何でもいい、要はオープンなものをみんなで分かち合うことによって創造性を高めイノベーションを生み出しようという考え方が生まれました。企業でも、古くから、例えば IBM がアーキテクチャーを開放したり、どこかの油田の会社が鉱山マップを解放して、油のありかを見つけてくれた人に報奨金を出したりという形でオープンイノベーションが生まれています。

一方で、オープンガバメントデータという形では、公共のオープンデータの方が先に世の中で進んでいき、科学にも影響を与え始めた。そうすると市民がデータに対する関心を高めていく。もともとサイエンスの時代から、サイエンスをどう公共社会とつなぐか、サイエンスコミュニケーション、アウトリーチという文脈で市民に提供するかという考え方はあると思うのですが、いずれにせよ市民がサイエンスに近づけるようになった。意識的にも、無意識的にも、近いところにサイエンスが来るようになったということになるのです。さらに Code for XX といった活動も生まれています。Code for Lib、Code for America、Code for Japan などが



(図 4)

あります。

このようなマップを書いて申し上げたかったのが、オープンアクセスからオープンサイエンスに学術情報流通の環境が広がる際に、様々なステークホルダーやトピックスが含まれるようになってきたということです。こういった文脈の中で、図書館の機関リポジトリがどういう役割を果たすかを考えてみるのは非常に重要なことではないでしょうか。

論文のオープンアクセスぐらいまでは、電子化や、電子化したものを改良する過程が漸次的に進んできました。ところが2010～2013年ぐらいになってくると、非連続な変化になりました。分かりやすく言えば、紙の論文を電子化したら、その購読スタイルはそのままに図書館と個人でIP、IDを管理してやりましょうとして開発をしていましたが、「オープンアクセスになったら購読管理は要らない。その方がみんなに届くではないか」というパラダイムシフトが起きました。これは見方によっては破壊的なイノベーションであり、もう少し大きな目で見れば、学術情報流通をリデザインするという文脈が含まれます。オープンサイエンスの概念が広がることで、どんどんその非連続の文脈が強くなってきています。その上で政策はどう作られるべきかということ、内閣の検討会などで議論させていただきました。

第4期科学基本計画（2011～2015）

政策的に見ると、実は第4期科学基本計画（2011～2015）から既にオープンサイエンスのことは書いてありました。第4期の科学基本計画が実際に策定されているのは2009年や2010年なのですが、「国は、大学や公的研究機関における機関リポジトリの構築を推進し、論文、観測、実験データ等の教育研究成果の電子化による体系的収集、保存やオープンアクセスを促進する。また、学協会が刊行する論文誌の電子化、国立国会図書館や大学図書館が保有する人文社会科学も含めた文献、資料の電子化及びオープンアクセスを推進する」とあり、当時の有川九大元総長がまとめられ、

これは非常に素晴らしいものだと思います。しかし、残念ながら、これを受けて行われた実際の活動が機関リポジトリの促進と学会への電子化支援にとどまってしまったということで、第5期ではどうより具体的な施策が打て、活動につながるかという話になっていません。

なぜ今研究データ？

なぜここで研究データが注目されているのか、おさらいします。やはり2013年ぐらいから大きな展開がありました。まず2013年6月にG8でオープンデータ憲章が提唱されました。ただし、中心はオープンガバメントデータであり、公的機関が生み出すデータをオープンにするのが大原則です。しかし、そこでG8科学大臣会合も行われ、オープンリサーチデータ、つまり研究成果のデータの方もオープンにしましょうというG8の文科大臣クラスの合意がなされています。日本では文科大臣の代わりに原山優子総合科学技術イノベーション会議の委員が出席されて、実は日本も合意しているのです。研究データをオープンにするというのは、政策的にはもう大枠で筋が示されている状態なのです。

また、その1カ月前の2013年5月、グローバルリサーチカウンシル（Global Research Council：GRC）が、オープンアクセスアクションプランを策定して、オープンアクセスをより推進することとしています。政策上、上のレベルからの圧力はどんどんかかっている状態です。あるいは、2014年10月、日本でもSTSフォーラムで、「第5回EU・日本科学政策フォーラムサイエンス2.0：変革する科学」がありました。サイエンス2.0が取り上げられ、これを機に内閣府のオープンサイエンスの検討会につながったと私は個人的に理解しています。

OAの潜在的便益（政策的観点）

では、オープンアクセスの潜在的便益とは何なのでしょうか。図書館の方々中心のときと政策では、少し

コンテキストが違います。Fact sheet: Open Access in Horizon 2020 では、ベネフィットとして、「研究を加速し成果を見つけやすくすることで研究開発投資の費用対効果を上げる」「同じ研究を繰り返すことを避け、研究開発コストを抑える」「境界領域や多領域にまたがる研究の機会を増やし、多分野の協調を促す」「研究成果の商業化を早く広い観点から行い、公共研究開発投資の効果を上げ、科学情報を基にした新しい産業を生み出す」ことが挙げられています。政策的な意味合いは、費用対効果や新しい産業振興という考え方が入ります。しかし一方で、やはり多分野の協調という形でサイエンスの発展も当然視野に入っているということも付記したいと思います。

公的資金の費用対効果 (Impact)

政策的には公的資金の費用対効果、インパクトの話がどうしても必要になります。これまでのサイエントメトリックス、ビブリオメトリックスを含むインパクト計量では、どうしても学術・科学的効果、およびそれに関連する教育的効果が語られることが多いですが、政策で見た場合は、経済的波及効果や産業振興への貢献も考慮されます。あるいは今だと、社会的波及効果、社会的な教育効果も考慮されます。ここをどう分類するかは本当に議論をすると難しいので、あまり突っ込まれてしまうと困るのですが、いずれにせよ、パブリックエンゲージメント、サイエンスアウトリーチのような考え方も含めて、同じ税金を使うのだったら、より広範なインパクトが高い方がいいという考え方に政策上はなっています。社会的波及効果の計量の実践は難しいのですが、今一番ホットなトピックです。

実際、Horizon 2020 にも先ほど申し上げたような多様な便益の紹介がありますし、アメリカであれば OSTP、NSF 等がブローダーインパクトという考え方を取っています。単なる科学・学術的インパクトだけではなく、相対的により広いインパクトを、研究費を取るときに書かなければいけない。書かなければ、申請すらできない状態になっています。あるいは、申請

する際の研究成果が、リサーチアーティクルという文言から、リサーチプロダクトという文言に全部変えられています。武田先生がご紹介されていたように、ソフトウェアやデータを含む、あらゆる研究成果をプロダクト（生産物）として再定義して、それを成果や実績として求めています。

イギリスにおいては、REF の今度のプログラムで、学術科学的効果以外のインパクトを 20% の割合で評価するという方向性で動いています。研究成果のインパクトは 65%、残り 15% は環境となっています。政策的にはこのような幅広い効果を見ようとしています。

科学技術外交としての OA

政策ではどうしても、オープンにするかクローズにするか考えるときに、国益という観点が入ります。公的資金で得られた研究成果を全て一律にオープンにして、一生懸命出した成果で、海外の研究者が利する、海外の産業が利するようなことがあっては国としてはまずいのです。知財を守るためのオープンアクセス政策です。オープンアクセスを議論すればするほど、実は何をクローズにすべきかの議論に収束するというのをいつも申し上げています。

あるいは、全ての分野で、研究者が一律なオープンアクセス化で劇的な便益を得るわけでも、もちろんありません。既存の強みはむしろそのまま生かす方向であるべきだし、これは、自由貿易と保護貿易の観点からもそうです。

国際的動向を踏まえたオープンサイエンスに関する検討会 (内閣府)

内閣府の検討会がこれから何をやるようとしているかということ、3月3日の「日本経済新聞」と3月6日の「科学新聞」に記事が出ているのですが、当面、公的資金を得た研究成果のうち、論文とその論文に付随する論文を立証するデータに関しては原則公開という方針を打ち出そうとしています(図5)。それ以外の研究データもできるだけ公開という流れで話が進もうと

しています。これは新聞記事なので、内閣府の趣旨のとおりではない表現もあるのですが、このようなことが既に紹介されている状況です。

オールサイエンス オープンサイエンスポリシーマップ

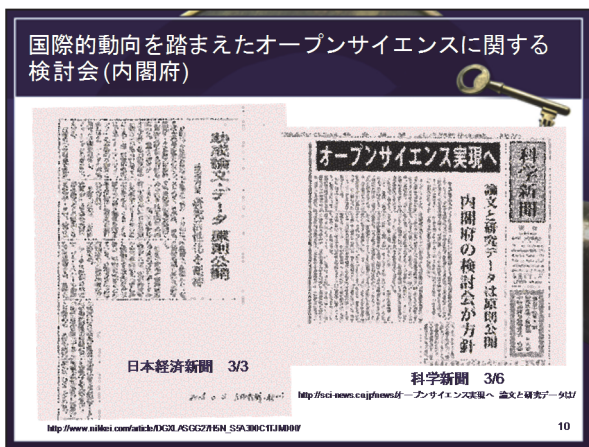
今までの話を「要はどうすることになるの?」というところでまとめたポリシーマップを作ってみました(図6)。まず、研究成果が出る前のデータ共有の話をするときに混乱するので、研究成果が出た後の話に限ります。そのときにまず必要なのが、「研究成果利活用ガイドライン」で、オープン、クローズの前に、まず利活用と言うようにしています。なぜならば、研究成果をクローズにして知財として活用し、産業を振興するルートがあってしかるべきだからです。オープンにして、敵対するステークホルダーが利するようなことがあってはならないのです。クローズにするなら今までどおり知財として活用し、研究データは登録しておいて、エンバーゴを設定します。すなわち一定期間後にオープンにします。いつでも必要に応じてオープンにできるようにして登録しておくということが、むしろ重要になってくるのです。

一方オープンにすると決めてしまえば、公開、共有を進める方針に従って、論文であれば Gold OA か Green OA を選択することになります。研究データであればリポジトリに登録します。それが主観型リポジトリなのか、機関リポジトリなのか、研究助成機関の

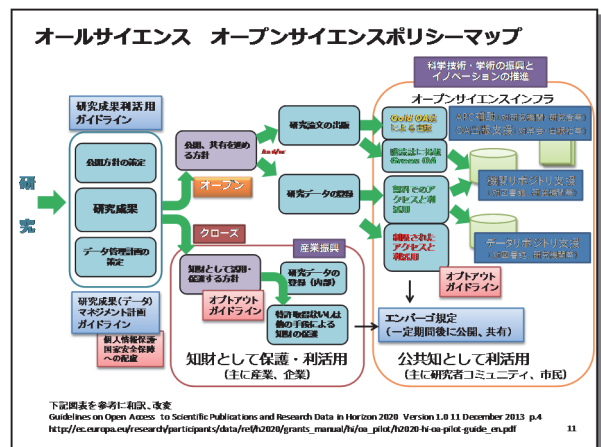
リポジトリなのかは、ケース・バイ・ケースですが、その中でも当然クローズの場合もあっていいと思います。

やはり、研究者がきちんと出した「マイバイビー」のデータですから、そのデータで自分たちが思う存分研究成果を出すまでは一定の猶予があってもいいと思います。ただし、公的資金を得られたものに関しても「飽きてしまった」ものは、エンバーゴの後にオープンにすることで、再利用によって新しいサイエンスが促進されます。科学技術の振興とイノベーションの推進という文脈です。

政策上の最大のポイントは、これからはオープンが原則でクローズにするときに理由をつくらなければいけなくなるかもしれないということです。研究者はそのように考え方を変えないといけなくなるかもしれません。まだここで確定としては申し上げられませんが、オープン化を進める際に今まではなぜオープンにするのかメリットを研究者に言わなければいけなかったのが、これからは、政策として公的資金を得るものに関しては、抜け道はあるとは言え、原則オープンにする、オープンにしないのなら、理由を付してオプトアウトするということになります。そこで「オプトアウトガイドライン」というものが必要になります。ここには個人情報保護、国家安全保障、倫理への配慮のようなことが書いてあります。何か理由がない限りはオープンにするということにこの図はなっています。



(図5)

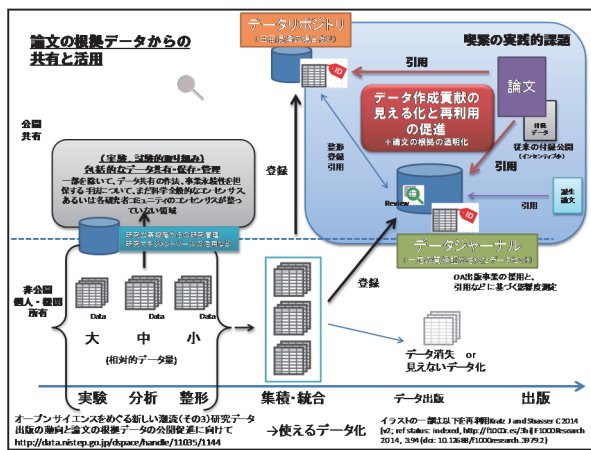


(図6)

論文の根拠データからの共有と活用

データジャーナルに関しては、やはり図書館の方を中心に注目すべきだと思っています(図7)。なぜならば、データジャーナルによって、商業出版社が頑張ってしまうからです。昔は論文に付録データが付いていて、これは今までどおりあるのですが、それがジャーナルに直接付いてなくてもデータリポジトリにもあればいいということで、機関リポジトリやサブジェクトリポジトリがあり、引用されます。ここでデータジャーナルが出てくることによって、研究データのお作法の質、すなわちこの研究データはどういう発想に基づいて作られ、どの研究者コミュニティの標準的なプロトコルを利用して出したものかというデータディスクリプターを加えることによって、一定の選別を行います。つまり、一定の信頼性をもとに再利用ができるのです。

安心して再利用ができるものを使って派生論文ができれば、ここでリユースファクターという再利用の指標に関する信頼性も高まります。と同時に、それを作るデータ作成者の貢献具合を見ることができます。仮定の話ですが、必ずしもデータ作成がうまい人が論文を書くのがうまいわけではないというときに、データ作成の貢献が一つの評価になる新しい可能性があります。今まで論文を書けている人たちからすれば、今のままで構わないのですが、こちらのデータ作成者の方に光が当たる可能性があります。このところで今、商業出版社がすでに様々な取り組みを進める中で、図



(図7)

書館の方々は機関リポジトリ等を通してどう対処していくかという考え方がやはり大事になってくるのではないかと思います。

話題提供：学術資料に関連した活動のご紹介

●堀井 図8は私のプロフィールです。私はもともと情報科学の出身で、少し前まで大学に勤めていました。その後、大学発ベンチャーの合同会社AMANEという小さな会社をつくり、今はそこで活動をしています。2014年10月に、一般社団法人学術資源リポジトリ協議会を設立しました。これから少し歴史資料の紹介をしますが、私自身は歴史、日本史の専門家ではありません、もちろん古文書も全く読めません。

一般社団法人 学術資源リポジトリ協議会

図9が学術資源リポジトリ協議会の概要です。法人成立が平成26年9月26日です。資料や古文書、教本などのいろいろな学術資源の情報資源化とリポジトリ化をしています。リポジトリ化というか、情報化、データ化するのですが、そういったこと自体を行う発想がない機関もあるので、そうしたところと一緒に、どのようにすれば資料の情報を広く横断的に、組織の枠や地域の枠を超えて共有することができるのかを考えていこうと、大学の研究者や大学博物館の方、理事の一人である山下先生、企業の人たちなど、学術に関わる人たちが、こういう団体をつくって地味に淡々肅々

自己紹介

氏名: 堀井 洋 (ほりい ひろし)

2002年6月 北陸先端科学技術大学院大学 情報科学研究科 博士後期課程修了 博士(情報科学)


2002年6月～2003年9月 文部科学省ITプログラム研究員

2003年10月～2011年9月 北陸先端科学技術大学院大学 知能科学研究科 複合システム論講座 助教

2009年4月 合同会社AMANE 設立 代表社員に就任

2014年10月 一般社団法人学術資源リポジトリ協議会を設立

専門分野: 情報システム学・人文情報学
所属学会: 情報地理学会 情報知識学会 アードコミュニケーション学会 加賀藩研究ネットワーク その他



(図8)

と活動をしているのが実情です。

多分、Web ページに幾つか出ているのですが、図 10 は、石川県立自然史資料館という県の施設にある、旧制第四高等学校（金沢大学の前身）の時代に作られた、虫の蛾が描かれている教育掛図のデータです。動いているのは NetCommons の上で、WEKO で実装します。WEKO をチューニングして、こういったものを公開しています。

現在公開中の学術資料情報

我々のところだけではなく、他の大学や自治体等とも連携しています（図 11）。教育掛図など、明治、大正の戦前の資料だけ抽出してきたのですが、こうしたいろいろな大学で、同じような資料が分散所蔵されているので、そういったものを集約して、どこに何があるかを共有することによって、何か新しい研究的な知見が得られるのではないかと期待して取り組んでいます。今はいろいろなプロジェクトが並行して進行しているので、どんどん資料が増えていきます。こういった活動はそれなりに認知されてきたのではないかと思います。最近思うところです。

Re*poN が目指しているもの

また、学術資源リポジトリ協議会という名前は長いので、Re*poN という愛称を付けています。何をやりたいかという、今までいろいろなところにあった所蔵資料やいろいろな資料を把握して資料目録を作り、

そういったものを電子化してメタデータを作ることです（図 12）。ただ、それだけではなく、例えば横断的に共有する場合には、その情報は本当に正しいか、きちんと研究者が検証したものなのか、あるいは、粒度と書いていますが、どういう単位で情報の固まりにしているのか、形式の調整もしています。また、固有 ID の付与、関連資料の整理を少しずつ進めています。「B.高付加価値メタデータ生成」とありますが、ただ単に資料集を作るのではなく、一手間加えた段階で公開したいと考えています。

さらに、組織間連携・渉外・契約は非常に難しいものがあるので、この部分についても手伝いたいと思っています。組織や機関との公開に関するさまざまな連携や合意、クリエイティブコモンズ準拠などいろいろありますが、権利処理の問題なども対処しなければ、最終的にはなかなか公開までたどり着きません。



(図 10)

一般社団法人 学術資源リポジトリ協議会

- ・ 名称：一般社団法人学術資源リポジトリ協議会
- ・ 主たる事務所：石川県能美市山田町口8番地
- ・ 法人成立の年月日：平成26年9月26日
- ・ 目的等
当法人は、広く学術資料を対象とした横断的な情報共有基盤の構築・整備とそれらに関わる人的なネットワークの形成を目的とし、次の事業を行う。
- ・ 事業内容
 - (1) 細粒・分野を横断した学術資料情報の共有のためのサービスの提供。
 - (2) 学術資源リポジトリの実現と普及に向けた活動。
 - (3) 学術情報及び関連する専門知識を活用した社会貢献活動の実施。
 - (4) その他、当法人の目的を達成するために必要な事業。

Re*poN
Academic Repository Network

第4回 SPARC Japan セミナー 2014 2014.03.09 11:00 会場: 金沢大学 学術情報センター

(図 9)

現在公開中の学術資料情報

明治～昭和期科学実験機器資料	
新潟大学所蔵：	2 0 点
神戸大学所蔵：	2 1 点
東京大学駒場博物館蔵：	2 2 点
石川県立自然史資料館蔵：	7 4 7 点
大阪教育大学附属図書館所蔵：	3 点
明治～昭和期教育掛図資料	
石川県立自然史資料館蔵：	1 2 6 点
大阪教育大学附属図書館所蔵：	6 1 点
奈良女子大学附属図書館蔵：	約 2 0 0 点

Re*poN
Academic Repository Network

第4回 SPARC Japan セミナー 2014 2014.03.09 11:00 会場: 金沢大学 学術情報センター

(図 11)

金沢大学 Virtual Museum Project

宣伝なのですが、同じような仕組みを使って、現在、金沢大学 Virtual Museum Project があります (図 13)。金沢大学は旧制第四高等学校で、いろいろな古い旧制高校時代の資料が残っている数少ない事例です。これ以外は、山下先生の京都大学が第三高等学校なので、三高と四高に資料が多く残っています。他にも東大の一高などもあるのですが、旧制高校同士の資料を比較することによって、またいろいろなことが分かっていくということもありますので、金沢大学は4~5年前から Virtual Museum Project を立ち上げて、資料館、あるいは図書館が持っている資料を公開しています。

本セミナーでの立場

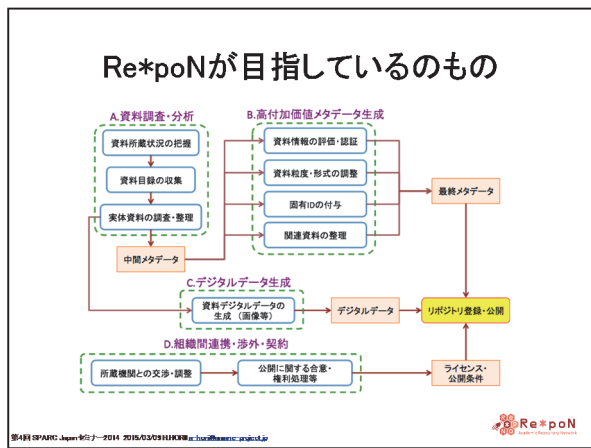
本日は、図 14 の左側の部分に関してご紹介します。林さんが非常にグローバルで政治的なお話をされたの

で、私は徹底的にミニマムな、自分の手の中のお話をします。

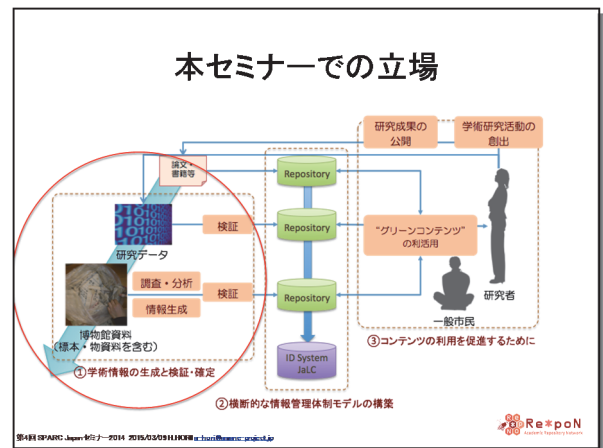
ある自治体に所蔵されている資料の例を挙げます。まずは、長持の中にたくさん大福帳などが詰まっているもので、恐らく北前船という日本海側を運んでいった船に関連した文書なのですが、全く手つかずの状態のものもあります。他には、天文関連の資料があります。天球儀ではないかと思うのですが、エアキャップが付いており、未整理であることが分かります。

図 15 は南山さんのスライドにも使っていたものです。これは私の手が少し写っています。こういう写真資料などもあります。

図 16 の写真は、山形県の恐らく庄内地方です。酒田や鶴岡あたりの戦前の様子を写した写真だと思われるのですが、まだ裏が取れていません。こういったものがたくさんあります。



(図 12)



(図 14)



(図 13)



(図 15)

研究データ生成の事例

私は今日、石川県から来て、石川県は加賀百万石の加賀藩なので、私が過去に行った研究の一つとして、「加賀藩先祖由緒井一類附帳」についてご紹介します(図17)。加賀藩は非常に大きな藩で、戸籍に当たる由緒帳というものがあります。それは戦災を免れて燃えませんでした。他に有名なところは徳島の蜂須賀藩があります。そちらは科研でもう公開されています。加賀藩の史料数は約1万1,760件です。過去に2代にわたって、私と一橋大学の林正治さんの2人で、科研で何とかデータベース化しました(図18)。これは結構大変でした。

図19は金沢の中心部に所在する玉川図書館近世史料館にある古文書の情報です。古文書を地元の研究者の方も含めた日本史研究者に解説していただいて、こういう形にしました。全部は抽出できなかったのです

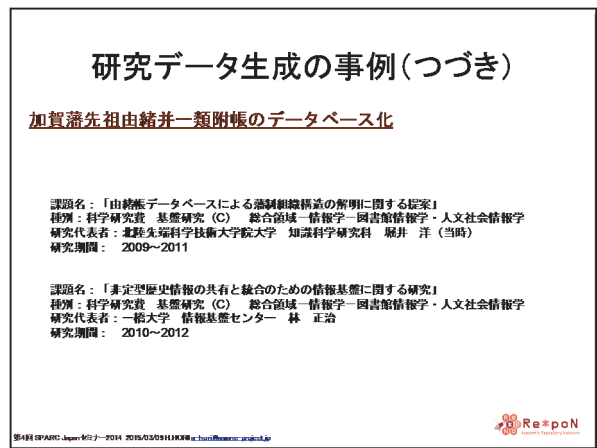


(図16)

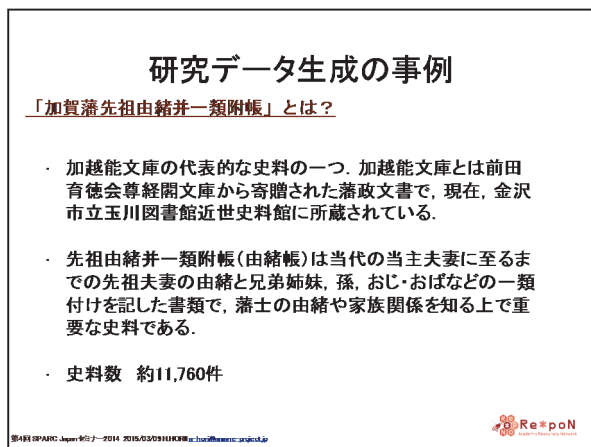
が、帙(ちつ)番号という箱の番号や名前、秩禄などのデータを作って、RDFにして出しています。

図20は、「加賀藩研究ネットワーク」という日本史の加賀藩を研究する人たちの研究コミュニティにおいて限定公開しているデータです。拡大したのですが、猪山左内とは、磯田道史先生の『武士の家計簿』の人で、引くときちんと載っていることが分かります。これ以外にも、加賀藩は、武士が非常に強かった藩だといわれているのですが、実際は、統計的に処理してみると、殿様の直々の家来の直臣ではなく、陪臣という家来の家来の人が多かったのです。半分近くが陪臣であったなど、いろいろ分かります。私はこれを一般公開したいと強く思っているのですが、今、課題にぶち当たっています。

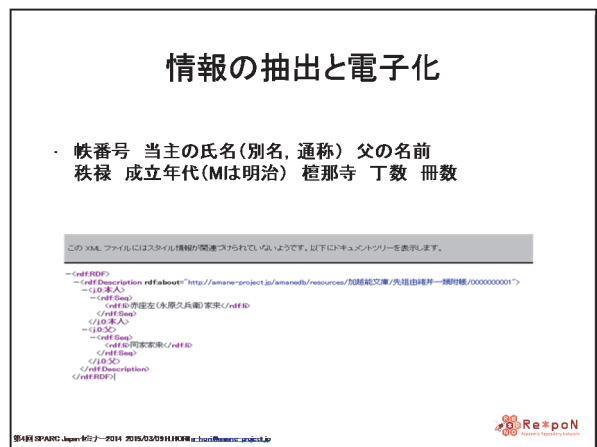
一つ目は、汎用的な公開研究データとしての精度・完成度がまだまだ低いことです。日本史研究者から、



(図18)



(図17)

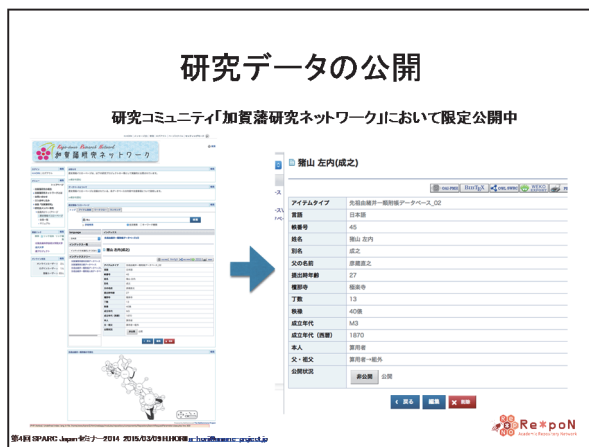


(図19)

「これを公開するのはまだ早いだろう」「まだ公開しないでくれ」など、いろいろな意見が出てきました。二つ目は、資料所蔵者・関係者の中での公開に関する承諾です。これは非常に難しい問題です。所蔵者の玉川図書館近世史料館との関係でまだ調整がついていません。科研でやるから使わせてくれと言ったのですが、それを広く一般公開するというのはまたちょっと違うということです。ちなみに、由緒帳は近世史料館で一番人気がある資料だそうです。多分、自分の先祖を調べに来るのだと思うのですが、そういうことがあるかないかは別として、今はこれが難しいです。三つ目は、一番現実的かもしれませんが、公開データを作って形を整える労力やコストが現状ないということです。

オープン化するために

データを作った研究者が、その研究データを直接他の研究者に使ってもらうというのが一番理想ではあるのですが、かなり厳しいと思います(図21)。私としては、間にデータの整形や監修、公開に向けた処理をしてくれる人がいるといいのではないかと考えています。このあたりも含めて現実的にデータを公開する。日本史は恐らく、物理や生物に比べて一番これが遅れています。遅れていると言っはいけないのですが、独特の世界があるのです。ただ、やはりこの世界は無視はしたくないし、無視できないと思うので、何とかうまくアプローチできればいいと考えています。



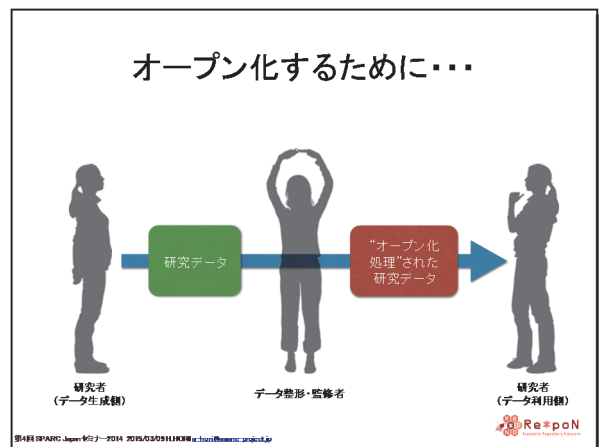
(図 20)

ディスカッション

●堀井 ここからはパネルディスカッションに入っていきます。初めに、講演のおさらいも含めて、パネラーの方々から一人ずつ、今回のテーマである、学術資源のオープン化の範囲を博物資料や研究データまで広げる目的や意義について、一言お願いします。

●武田 オープン化については、外側から来るプレッシャーや状況の変化は今後もあるかも知れませんが、私の興味はむしろそれに図書館なり機関リポジトリがどういふふうに対応するか、それに目的・意義を見つけられるかということです。つまり、一般論として学術資源のオープン化うんぬんは、林さんの話にあったようにもう我々の外で動いていて、これは僕の中から見ればもう止まることはないでしょう。今、黙っているとオープン化になる、白と黒が逆転する段階にいます。5年ぐらいたつと、「昔はクロードがデフォルトだったよね」と言うと、「え？ 何で？」と逆に聞かれる時代になるでしょう。

機関リポジトリが取り込むかどうかというのは、データの共有や公開が研究の重要な一プロセスになっている段階において、自分たちが入れるかどうかということです。機関リポジトリが単なるシステムの手助け屋さんでは困ります。質疑応答でフロアからコメントがあったように、「そこに何らかの見識を出してほしい」というのは確かにあると思います。それは今ない



(図 21)

わけです。各分野でいい意味で勝手気ままにやっている段階です。ただ、それが業績として認められるということは、それほど勝手気ままでは許されなくなるので、そのときに、少しはそこに対して貢献ができるか。機関リポジトリが学術資源のオープン化にコミットするための目的・意義を設定できるかというのではないかと思います。

●堀井 先ほど、学術分野の分野依存性が非常に強いという話がありましたよね。それは、これからほとんど弱まっていく方向に行くとお考えですか。

●武田 いえ、それは変わらないというか、むしろ全体で見れば拡大するはずですよ。ただ、独自性はどんどん広がるけれども、共通部分を見つけることはできるでしょう。今は分野を超えて共通部分を見つける努力すらしていないですから。

●山下 リポジトリが扱う対象を広げる意義ですね？博物館の方で何年も前に、データベースをたくさん作って公開するという流れがあったのですが、それはそれなりに意味はあったにしても、あまりうまくいかなかったのではないかと思います。やはり論文、電子ジャーナルみたいなものが、これだけ揃ってくるようになって、今、リポジトリであれ、博物館であれ、もう一度自分たちの標本や資料を出すことで、そういう論文のようなわかりやすい研究成果のものときちんとつながるタイミングなのではないかなという気がしています。図書館にも特殊コレクションや博物館のような資料もありますから、それが今、そこにまで対象を広げてやることの意味だと思います。

●南山 図書館の文脈でお話すると、出版という行為自体が、そもそも整理して、出版して広げていくという行為ですから、それを保存して公開している図書館の仕事というのは、もともと本質的にオープン化に関わるものではないかと思っています。それが電子媒

体になっても、役割は変わるものではなく、本質的な仕事として捉えていくべきではないかということがオープン化の意義です。

●林 機関リポジトリを作るメタレベルの目的は、ただ単にアーカイビングするのではなく、機関リポジトリというシステムを軸に所属の研究者の研究活動のプレゼンスを上げることです。学術情報流通に関して日本のベンダーがもっと活躍し、産業振興につながるという観点を図書館の方々も持って、積極的に図書館を使った機関リポジトリを発展させると、日本の産業も良くなる可能性があります。そうしてお金回りが良くなると、税金が増えて、税金が増えると公的な研究費が増えて、風が吹けば桶屋がもうかるということになります。学術情報流通に関連する産業の振興なくして研究の振興もないということが、政策の方に携わるとよく分かるようになってきたのです。図書館の活動だけの局所的なところで最適化を図るのは避けた方がいいし、既存の枠を飛び出ればもっといろいろな可能性があるのではないかと考えています。

先ほどデータジャーナルのところでも少し時間が押していたので申し上げられなかったのですが、データジャーナルのクオリティコントロールは、データの中身の品質を見るのではなく、データの記述様式の品質を見ます。これは図書館が目録管理をするときの行動原理に非常に近いので、データジャーナルにおいては図書館の人々の知見が大いに役立つ余地があります。リポジトリの上に、例えばですけれども、データジャーナルを作って、その活動プロトコルが広がっていくということもあっていいのではないかと思います。

●堀井 ありがとうございます。今回は決起集会ということでしたので、結局どうすればこの状況を変化させてオープン化に向けて進むことができるのかという部分が一番重要だと思うのですが、そのあたりについてどうですか。

●**南山** また図書館の文脈でお話ししますが、先ほど林さんのお話を聞いて一番思ったのは、どこまでをオープンにして、どこまでをクローズにしなければいけないという線引きを早くしないと、オープン化も進まないのではないかとということです。ですから、まず図書館でリポジトリを扱うのであれば、どこまでを対象として、マンドートでも推奨レベルでもいいですが、ここまでは出してほしい、ここまでは先生方にお任せする、あるいは、業績として認められるのであれば、こういうふうにするといったポリシーの策定を、図書館だけでなく、ひな形を作って先生方とご相談するといった取り組みをできるだけ早く行うことが一つだと思います。

もう一つは、そもそも図書館員が主題を理解できれば話は早いと思うのですが、それはなかなか難しいところもあるので、今の人材で対応できること、例えば博物館の例でもありましたが、大学院生の知識を使って一緒に協力して目録を作るとか、URAの人と協力するとか、ケース・バイ・ケースになるのですが、それを各人が当事者意識を持ってやっていく必要があるのではないかと思います。

●**堀井** 林さん、そのあたりは国内外の状況も含めていかがでしょうか。

●**林** 南山さんの発言に一つ付け加えると、やはりデータに関してはデータマネジメントプランへの関与の仕方も大きなキーポイントです。サブジェクトライブラリアン化への一歩と見ることも可能だと思っています。すなわち、研究データマネジメントプランを作ろうと思うと、研究を計画する段階から図書館がコミットして、そもそもその研究は何のためにやるのか、それによって得られる成果、アウトプットとプロダクトは何なのかという議論から入っていくことになります。そのトレーニングを図書館全員が行う必要はないかもしれませんが、入り込むことは一つの活動案になり得るのではないかと考えています。一方、もちろんポ

リシーの方は、むしろ、主にリサーチ・アドミニストレーターと協働して策定していく。この二つは確かにアクションプランとして当然の流れなのではないかと思っています。

●**堀井** 我々だけではなく、いろいろな人たちから何をすべきかについて意見を伺いたいのですが、山下先生のとくも武田先生のとくも質問時間が取れなかったもので、それも併せて、何か会場の方からご意見があれば、ぜひお話しただけでないでしょうか。

●**フロア1** 九州大学附属図書館の職員です。山下先生への質問です。学術資源アーカイブの取り組みにすごく興味があります。今、九州大学内で附属図書館と博物館と文書館と資料館で、まさに同じような取り組みをしようとしていて、相互に持っているデータベースの資料を横断的に閲覧できるような取り組みをしようとしています。

そこで問題になっているのは、博物館でたくさん資料は持っているのですが、その資料をもともと使われた方、作られた方がご退官されていたりして連絡が取れないことです。そうすると、これを再び利用している、公開していいというライセンスが分からないので、基本的には公開できないものがたくさんあります。ライセンス不明な資料を、これからどういうふうに取り扱って公開していけばいいのかと悩んでいます。もし、ライセンス不明な資料に関して、こうしたらいいのではないかという経験や取り組みがあったら聞かせていただければと思います。

●**山下** 僕以外の先生の方が詳しいのではないかとと思うのですが、研究資源アーカイブは割と現実的な対応をしています。どうせ全部一度には公開できません。公募するときに公開できる資料の評価が多少高くなるのです。自動的に権利が処理できなさそうなものは、公開しないけれど、取りあえずデジタル化するなり、手当てをするなりという道をとるのですが、基本的に

は公開できるものを優先的に選んでいくという逃げ道をつくっています。

●林 日本の学会の論文電子化、アーカイブ化のときに同じことが起きました。著者が亡くなってしまって連絡がつかなかったのです。杓子定規には文化庁裁定に持ち込まないといけません。要するに、供託金を預けて、新聞広告か何かに出して、「著作権をこちらで管理します」と周知する方法です。お勧めはしませんが、その当時、学会で何をやったかという、ネット上で「これこれの論文をうちの学会、出版社のアーカイブとしてオープンに公開します。文句がある人はこの日までにここまで言ってきてください」と言って、そのまま公開してしまいました。なぜこのような乱暴なことができたかという、学術資源においては、権利関係に絡む訴訟になるほどの案件になることはほとんどないからです。むしろ、成果を広げた方が喜ばれるという構造があるからできたのです。それがだんだん博物学資料になってくると、微妙にうるさいことを言う人が出てきそうな感じなので、とてもお勧めできるものではありませんが、一応ご参考までに。

●堀井 私も関わった例ですと、最初はやはりデータとしてはあまり細かい、粗々としたものだけ公開して、どんどん時間を置くにつれて……。文化庁裁定はかなりハードルが高いです。現状では現実的ではない場合が多いかと思います。

●フロア2 そのようなことを言っても、誰も権利を主張していないのだから、学術の目的のために公開して何が問題なのですか。

●フロア1 できれば大学としても公開したいのですが、もし何かトラブルがあったりしたときに。例えば、「私はそれを公開していいとは言っていない」など。先ほど南山さんが講演で言われたように、たくさん苦勞して得られたデータなので、勝手に使われる

ことを不快に思う人たちもいて、それを大学が連絡が取れなかったから許可なしにやったという話になると、データを作られた本人の方々には不快に思われる方もいるかもしれないと思いました。

●フロア2 現れたところで、誠心誠意話し合えばいいことではないですか。僕はもう一つ伺いたいです。堀井さんから、日本史関係の資料の公開うんぬんという話がありましたが、そのようなものを公開する必要がそもそもあるのですか。実物資料が読めれば、そこにやってきて読んでいただく方が、より良い資料の公開ではないですか。実物については何の問題もないわけでしょう？ 所有権というわけの分からないことを主張される方は、実物資料の方がこだわりが多いようなのですが、図書館でも博物館でも、学術目的のために、来たら見せてあげるといのは、今までもずっとやっていることですから、特に問題がないので、何もそれをインターネットでオープンに公開する必要はないでしょうし、所在の情報さえ教えてあげれば問題ないのではないですか。そのためにはDOI、デジタルなオブジェクトアイデンティファイヤーがあればいいだけのことなので、ものがデジタルである必要は全然ないわけでしょう。

●堀井 公開するのは全部ではなくて、少なくとも、外から検索できるのは、どこに何があるかという存在だけです。

●フロア2 だとすれば、何の問題もないですよね。メタデータでは、権利関係の問題はほとんど生じるはずがないのですから。

●堀井 それも含めて今後、史料館側と話したいと思います。

●武田 日本史資料が少ないと言われたのですが、例えば、外国にも日本の文献を研究している人はいるの

で、そういう人にとって博物館に来なければアクセスできないというのは、もはや通じないわけです。実際、海外の博物館ではそういう資料は全部電子化してアクセスできるようにしているのに、日本は特にそういう必要がないというのは、研究のスタンダードに合わないと言われるだけですよね。

●**フロア2** ですが、そのような日本研究をしている人は非常にわずかで、多くの人が日本に対するアクセスをたくさん持っている現状においては、それほど頑張らなくてもいいのではないですか。もちろん外交的な観点から言って、海外諸国が研究資料を公開しているにもかかわらず、日本は隠しているように見えるという問題は解決しなければいけないと思いますが、そもそも公開しなくてはいけないという議論とは別だと思えます。研究上の目的からすれば、いつでもアクセスできるのであれば十分なのではないですか。

●**武田** 多分、十分ではないと思います。

●**フロア2** 十分ではないとすれば、類似の多くのものを比較することによって、初めて可能になるような研究の種類があるのだから、そういうものためにはデジタル化すること自体が必要で、デジタル化されたなら、一定のライセンスの下に、利用可能な状態になっているというのは当然だというご主張なら分かるのです。そうではなく、海外もやっているから日本でもやらなくてはというのは変な議論だと思います。世界がそうやっているからといって、それが合理的な理由になるとは思わないのですが、合理的だからやるべきだとおっしゃるなら分かるのですが、他のところでやっているからやるというのは変だと思います。

●**武田** 研究の世界のトレンドが今そういう状況になっている中で、日本だけが「日本という特殊事情があるからできません」と言うのは通じないということです。

●**堀井** 最後に、今後に向けた決意表明ということで、オープン化に向けて今後どういうふう活動していくか、もう実際に活動されている方もいらっしゃいますが、何か一言あれば頂けないでしょうか。武田先生から。

●**武田** 研究の世界が大きく変わり、デジタル化の研究になった中で、さらにいろいろなものが世界で共有されるので、そこはもう世界と協調的にやらなければいけません。一方でそれは競争でもあります。先行している人たちが有利になるという結果的な仕組みがあるので、そのためにはフォロワーにならずに、我々も協調と競争の世界に早く入れるようになりたいと考えています。私も個人的にそういう努力はしますが、特に海外ではデータリポジトリなども図書館側になっているケースも多いので、その辺に関しては日本ではどのような取り組みをするかというのを、ぜひ皆さんと一緒に考えていきたいと思っています。

●**山下** 図書館ではないので立場は違うかもしれませんが、私自身は、古い資料のアーカイブ、いわゆる使い終わった、現用ではない研究データのアーカイブにも関心があります。過去の研究データだったものを使ってオープン化していくことで、現在研究が進行している研究データのオープン化についても何か分かることがあるのではないかということです。そのようにアプローチしていきたいと思っています。

●**南山** 私は、これだけ研究データ、研究データといわれている時代で、これはそもそも大学図書館の仕事だと思っているのです。大学図書館の仕事なのに、電子だから対応できないと言ったら、「では図書館は要るのですか」という話になりかねないという危惧をずっと持っています。ですから、保存、流通を担う図書館としては、初めの段階できちんとこの仕組みの策定に関わり、持続可能な仕組みをつくるために図書館の意見を反映させていくべきではないかとずっと思っ

います。

●林 私は今日の活動も含めて、研究者を中心とした皆さまの自発的な活動をなるべく政策に反映させる、あるいは行政官にこの状況を理解してもらい翻訳者のような立場に結果的になっているので、そうした活動を通じてより実効的な政策が打てるように、そして元は実験室にいた科学者としては、何か新しいわくわくする研究活動が生まれるような場を作れるように頑張っていこうと思っています。

●堀井 このオープン化が進むことによって、古い資料が見直され、整理も進みます。データへのアクセスだけではなく、データから実物資料へのアクセスが確保されることが必要だと思いますし、それに携わる学術専門人材も、新たなポジションが生まれることを期待しています。

国際化という観点では、先ほどお見せした加賀藩研究ネットワークは、たかだか70人ぐらいしかいない研究コミュニティなのですが、日本ではなく現地にいる、フランス人が1人、アメリカ人が1人、ニュージーランド人が1人会員になって、加賀藩というローカルな日本史の研究をしています。これを一つの例として、新しい研究ができるような環境を将来的につくっていければと思います。パネルディスカッションは以上で終了します。どうもありがとうございました。