

第1回 SPARC Japan セミナー2017

「図書館員と研究者の新たな関係: 研究データの管理と流通から考える」

研究者にとってのデータの意味と 大学におけるデータ管理への期待

倉田 敬子

(慶應義塾大学文学部)

講演要旨



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



倉田 敬子

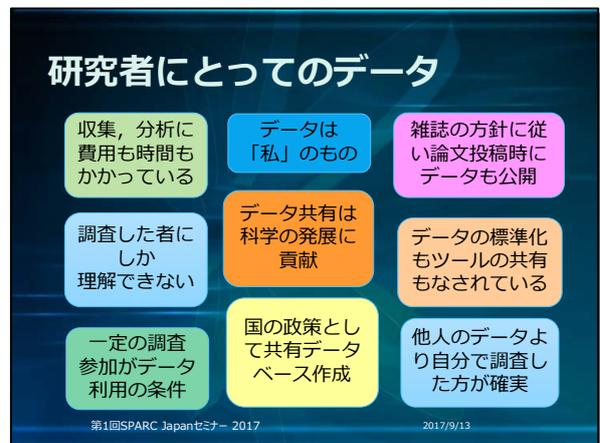
1987年慶應義塾大学文学研究科博士課程修了。1988年慶應義塾大学文学部図書館・情報学科助手、1993年同助教授、2001年から文学部図書館・情報学専攻教授。専門は学術コミュニケーション、特に情報メディアのデジタル化、オープンアクセス、研究データに関心を持っている。著書に『学術情報流通とオープンアクセス』（勁草書房）など。2008年度日本図書館情報学会賞授賞、2010年科学技術への顕著な貢献2010（ナイスステップな研究者）選定。

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

に見せることに、何となく抵抗があるのです。ただ、これは分野によって全く異なります。例えばヒトゲノムを中心とする生物医学分野では、「そんなものは出

1. データ共有の阻害要因と推進要因

研究者がデータの共有や公開、もしくは研究データそのものに対して持っている意識は大変多様です（図1）。もちろんデータ共有は科学の発展に貢献すると主張し、自らも実践している方もいますが、一方で研究者には「そう言われてもデータは私のものだよね」という感覚がやはりどこかにあります。自分が関わり、収集や分析に多大な費用も労力もかけたデータ、それを他人



(図1)

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

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

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

彼らの行動をある意味では制約して、推進・阻

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

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

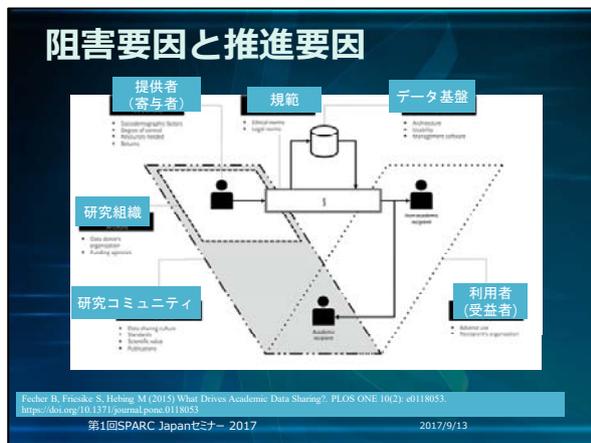
2. 研究者の意識・行動のモデル化

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

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

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

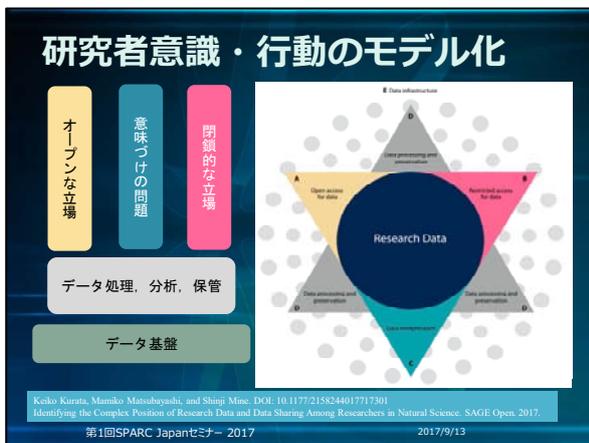
ここからは私達の方で解釈したものですが、それを五つのグループに分けられるのではないかと考えました。図 3 の右側が論文に出した図です。「Research Data」が真ん中にあり、それを囲っているイエローと



(図 2)

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

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



(図3)

3. データとは何なのか？

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

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

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

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



(図4)

ロセスとは異なるのです。

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

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

そう考えると、研究データを管理するには当然のことながら研究プロセスへの理解が必要です。研究者がどのような段階で何を求めている、だから、今このデータなのだということを理解した上で、データの管理

を考えなければいけません。データの管理は、研究支援全体の文脈の中で考えなければいけません。データだけを「このような形できちんとしてください」と言っても、それはやはりなかなか難しいわけです。なぜそのデータをそのような形で管理していかなくてはいいけないのか、例えばデータ管理計画をなぜこんなときに出さなければいけないのかというのは、研究プロセス全体とデータの管理全体の両方をよく理解しないと、研究者と管理側の両方の考え、認識を一致させることはできないだろうということです。

4.大学・研究機関における支援の現状

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

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

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

なぜ複雑、多様になるのか

- 「データ」と研究の不可分な関係
- 研究プロセス中心
- データが「研究実施」段階でしか明示されていない
- データの段階がうまく表示できていない

Source: <http://books.jiscinfonet.ac.uk/vse-lifecycle/index.html>
第1回SPARC Japanセミナー 2017 2017/9/13

(図5)

大学における研究データの管理、保管、公開

- 2016年3月実施の調査結果紹介
- 情報管理.2017, vol.60, no.2, p.119-127
- [対象]大学・研究機関494→151回答(30.6%)
- [設問]
- ①オープン化への認識 ②OA化状況
- ③研究データ保存ガイドライン制定
- DMPの現状・支援
- 研究データ用プラットフォーム整備状況
- データ公開方策の今後の可能性

第1回SPARC Japanセミナー 2017 2017/9/13

(図6)

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

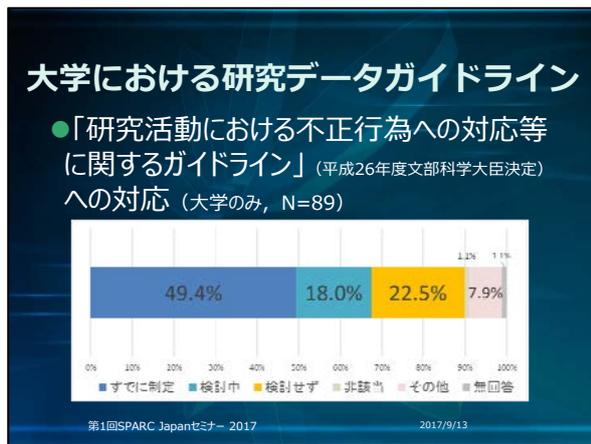
幾つか結果をご紹介します。「研究活動における不正行為への対応等に関するガイドライン」への対応については、大学が対応するものなので、大学のみデータを出力しています (図 7)。

「すでに制定」が 49.4%と約半分、「検討中」が 18.0%で、検討中も含めると、7 割が基本的にはガイドライン自体は制定しているということです。ただし、慶應義塾大学をはじめとする幾つかのガイドラインを見ると、基本的には研究者に丸投げなのです。保管はしろ、けれども大学は特に何もしない、それでも一応

ガイドラインはつくったということになります。最低限の認識はあるし、研究不正に対処する形での対応はしている、でも DMP については全く知らず、DMP の提出を義務付けているという回答は一つもありませんでした (図 8)。DMP の作成を推奨しているところもごくわずかです。

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

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



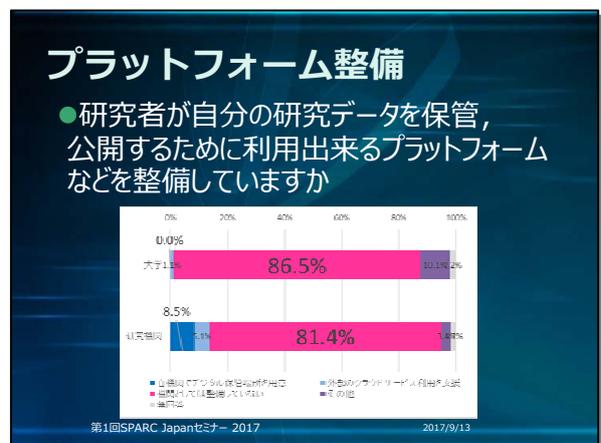
(図 7)



(図 9)



(図 8)



(図 10)

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

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

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

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

5. 図書館に何ができるのか

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

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

新しい役割の可能性としては、「①Embedded Librarian

ian」と「②学術情報支援部門の構築」という二つの方向性があるのではないかと考えています。

5-1. Embedded Librarian

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

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

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

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

かった要因があります。

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

5-2. 学術情報支援部門の構築

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

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

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

6. まとめ

このような大学の現状を考えた上で、それでもなお

かつオープンサイエンス、オープンデータを中心とした学術研究支援に図書館が乗り出すのだとするならば、一体何ができるのでしょうか。

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

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

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

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

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

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

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

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

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

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

と思うのです。その意味で、今、両者の関係をうまくつくるのが難しくなっています。

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

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

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

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

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

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

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

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

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

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

方はいらっしゃいますでしょうか。いない模様です。すみません、NIIの図書室にも届きませんでした。

●倉田 申し訳ありません。どこかに行ってしまったのだと思います。

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

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

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

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

大学における研究データの
管理, 保管, 公開

●2016年3月実施の調査結果紹介
情報管理, 2017, vol.60, no.2, p.119-127
[対象]大学・研究機関494→151回答(30.6%)
[設問]
①オープン化への認識 ②OA化状況
③研究データ保存ガイドライン制定
DMPの現状・支援
研究データ用プラットフォーム整備状況
データ公開方策の今後の可能性

第1回SPARC Japanセミナー 2017 2017/9/13

(図6)

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

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

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

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

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

ういう感じで、お答えになっていますでしょうか。

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

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

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