

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？」

図書館によるデータ管理への道筋

南山 泰之

(国立極地研究所)

講演要旨

近年、オープンサイエンスをキーワードに、データのオープン化の機運が高まっており、海外では図書館が積極的にサポート体制を構築しつつある。一方で日本国内では、データそのものの管理・保存に対する取り組みは主として研究者コミュニティや博物館等によって進められてきており、国内の図書館員が積極的に関わってきたとは言い難い現状がある。本報告では、データ管理を巡る国際情勢やこれまでの図書館における取り組みを振り返りつつ、今後の「グリーンコンテンツ」拡大のために、図書館はどのような役割を果たしていくべきか、研究者や助成機関、博物館等のコミュニティとどのように連携していくべきか、明日から何を始めるべきなのか、を議論したい。



南山 泰之

2005年より国立極地研究所情報図書室に勤務。2007年～2008年、第49次日本南極地域観測隊に参加。その後、東京大学駒場図書館（2011年～2014年）を経て現職に戻る。2014年8月より機関リポジトリ推進委員会コンテンツワーキンググループメンバー。

私は2005年から国立極地研究所に勤めはじめ、その約2年後にいきなり観測隊に放り込まれ、その後、東京大学教養学部に勤務し、昨年、また極地研に戻ってきたという経歴です。8月から、機関リポジトリ推進委員会のコンテンツワーキンググループ（WG）に所属して、データ関係の検討などを行っています。

観測隊では、オーロラの観測を手伝ったり、野外に行きペンギンを勘定したりしていました。49次隊のときは、「しらせ」が最後の年でした。今は「新しらせ」という新しいものになってしまいましたが、自衛隊の方々と一緒に行き、盛り上がり楽しかったのを覚えています。

本セミナーの位置付け

そもそものグリーンコンテンツとは、グリーンジャーナルに掲載された論文を機関リポジトリに搭載したもので、研究データや博物館資料、大学が出版局であるような紀要は対象外です。しかし、グリーンは研究機関のイメージ、ゴールドは出版社のイメージを持ち、グリーンの若葉が青々と育つイメージはコンテンツの拡大を連想させるため、本セミナーでは、グリーンコンテンツを「学術研究機関が機関リポジトリ等において、公開・発信する学術コンテンツ」として再定義します。

次に、オープンアクセスとは何かを簡単に振り返ります。「Budapest Open Access Initiative（2002年2月）」では、オープンアクセスを、無料かつ誰でも利用でき、

著作者の明示と、正当な評価と引用を求めるものと定義しています。2012年には、「次の10年に向けた提言」も出ており、クリエイティブ・コモンズ・ライセンスでCC-BYを推奨するなど、より具体的な提言もされていますが、まずはこの定義からになります。

その定義を受けてかどうか分かりませんが、オープンアクセスは全世界レベルで広がっています(図1)。上側の緑の点はリポジトリがどのように広がっていったか、下側の点はジャーナルレベルでオープンアクセスがどのように広がっていったかを表しています。

オープンアクセスが非常に広がってきたところですが、昨今、新たなキーワードとして「オープンサイエンス」という単語がよく聞かれます。オープンサイエンスは、ICTの発達によって生まれたアプローチであって、新たな知の創造プロセスであるという言い方をされます。オープンアクセスは、従来の論文のオープン化がメインで進んできたものですが、オープンサイエンスは、それにとどまらず、そもそもの根拠となるデータや、研究プロセスの公開などを含めて公開していこうという概念です。この概念の広がりや、異なる分野へのイノベーションの引き金や、利活用の促進、公共への還元、さらには説明責任といった社会的責任を果たすことにもつながってきます。

オープンサイエンスに期待されることは、知識の循環、再利用によるイノベーション、つまり、他業界との連携促進です。従って、本セミナーは、博物館資料や研究データなどのデータの連携、管理手法の検討を



(図1)

目的とするオープンサイエンス活動の一環として位置付けられます。

なぜデータが重要なのか？

さて、そもそもデータとは何なのでしょう。博物館におけるデータは、資料画像データやメタデータなどです。研究・観測データは、数値データやテキストレコード、イメージ・ビジュアルデータ、メタデータなどです。つまり、さまざまな研究の根拠や成果になるようなものです。

データの取得方法を、極地研を例に説明します。高度2,000mのセールロンダーネ山地まで行って、1週間かけてこつこつ石を掘り続けたり、湖に行って、湖の中に機械を沈めてデータを取得したり、ペンギンの営巣地に行って、何匹いるかを数えたりします。これらは非常に大変です。しかも、実験地にいられる時間は少ないので、ペンギンの数が多い場合は、写真を撮って後で見ながら数えます。お手伝いをしましたが、5分ぐらいで「もう勘弁してください」と言いました。

その後、まだ根気の要る作業は続きます。データの分析を一つ一つしていくのです。申し上げたいのは、観測データや研究データというのは、これだけやってやっと取得できるということです。ですから、何かメリットがなければ、オープンにしると言っても、簡単にできる代物ではないということをまず実感として共有いただきたいです。

図2は、オープンサイエンスによる科学研究の在り方を図で表したものです。研究の根拠や成果などのデータをもっと効果的に活用したいという試みを前面に押し出すものとして、データ中心科学(第4の科学)という在り方が提唱されています。

データ駆動型とは、従来の研究者一人一人がデータを持ってデータベースに保存していた形式から、データ利活用のプラットフォームを1個設けて、ここから研究者にデータが行き、そのデータを基に論文や研究データなどを作り、それをまたプラットフォームに戻すという、データを中心にした研究の流れがあり、プ

プラットフォームに情報が蓄積されていくイメージになります(図3)。この考え方では、保存・公開を担うプラットフォームが非常に大切です。

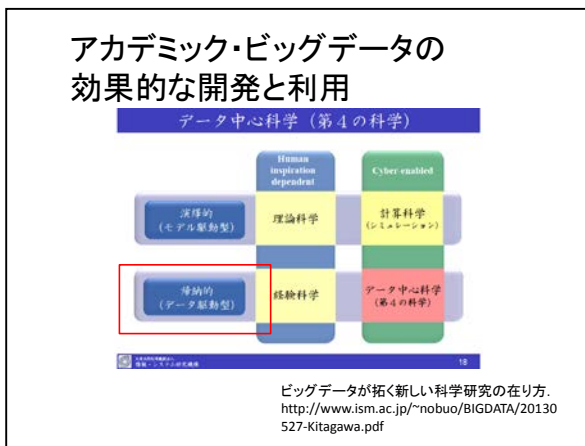
一方では、研究データのガバナンスという切り口からも、データの重要性が叫ばれています(図4)。一人一人がデータを管理していくという手法では、人事異動や担当者の交代によって、データが散逸する事態を招きやすいです。そうなると、最終的に困るのは責任者になります。

研究の根拠や成果のデータを保存・公開する意義には、オープンサイエンスの前提や共有による研究の効率化というメリットがあります。国際プロジェクトなどではデータの受け渡しは手間になるからです。また、管理者側としては研究不正への対応も挙げられます。研究の透明性や再現可能性といったキーワードで語られる文脈です。データの散逸を防ぐということです。

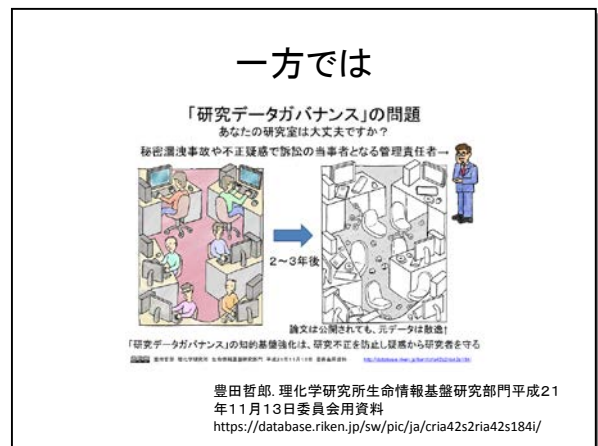
海外の動向

海外でこのような重要性を認識しているところは、積極的に公開支援を始めています(図5)。イギリスの大きなところでは、Digital Curation Centre (DCC)、DataCite、Research Data Alliance (RDA) や、オーストラリアの Australian National Data Service (ANDS) などが公開を支援しています。大学図書館の取り組みとしては、エジンバラ大学やパデュー大学で、プラットフォームの準備やガイダンスによる情報提供を始めています。

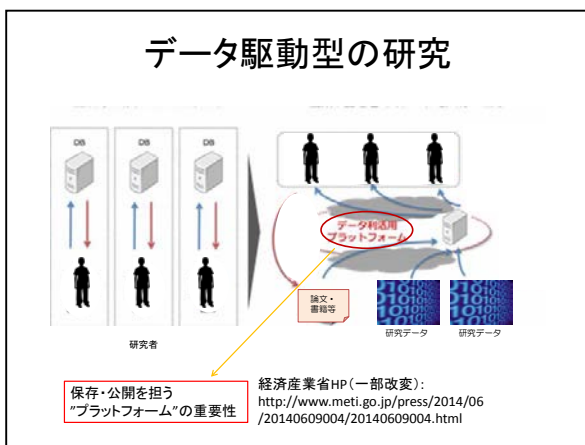
データを取得するのは非常に大変なので、何かしらメリットがなければ公開には行き着きません。従って、業績として認められるのであれば、データ公開の大きなインセンティブになるため、そのようなことも検討されています(図6)。



(図2)



(図4)



(図3)



(図5)

国内での取り組み

国内では、大学博物館と連携したデータリポジトリ普及の実践や、管理組織や体制の整備ということで、研究データへの DOI 付与実験プロジェクトなどの取り組みも始まっています。

オープンデータというくくりで言うと、東寺百合文書(とうじひやくごうもんじょ)を京都府立総合資料館の Web で CC ライセンス、CC-BY で公開したオープン化の試みなども話題になりました。

データ管理をめぐる国内動向は、2014 年 12 月から、内閣府で、「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」が開かれています。また、2014 年 12 月、国立国会図書館でも、「科学技術情報整備審議会/第四期科学技術情報整備基本計画策定に向けた基本方針検討部会」が開かれました。さらに、2015 年 3 月 4 日、日本学術会議学術フォーラム「科学を変えるデータジャーナル—科学技術データの共有・再利用の新たなプラットフォーム構築に向けて—」も開催され、ここ 3~4 カ月、動きが激しくなっている印象を受けています。

図書館が果たすべき役割の検討

その中で、日本の図書館は何をやっているのでしょうか(図 7)。蚊帳の外ではなく、これから頑張ろうという話をします。

機関リポジトリでのデータの取り扱い現状を見ると、データやデータセットは 3.6%で、さらにほとんどが

千葉大のものというように、取り扱いは体系的ではありませんが、やろうという試みは進んでいます(図 8)。

NII と大学図書館で機関リポジトリ推進委員会を立ち上げ、コンテンツ WG でデータの取り扱いを検討しています。昨年 8 月に WG が発足し、11 月に図書館総合展でエジンバラ大学のシュワート・ルイスさんをお呼びして、具体的なお話を聞きました。今年 2 月に入ってから、国際 WG と共同で、データ関係のカンファレンスである IDCC (International Digital Curation Conference) に参加しています。

私が所属している研究データの動向把握および試行収集班(通称:データ班)のタスクは、なぜ図書館が研究データを扱うのかを明らかにすることと、主に海外の研究データ管理の動向調査です。海外ではどうしているのかという疑問を持って、図書館総

で、日本の図書館は?

蚊帳の外?

(図 7)

データの業績化

NSF(米国立科学財団)
2013年1月から業績としてデータを記載可能に
"Publication" → "Products"

リユースファクター(提案)
データやコードの引用によるインパクトを測定

転載元:
「オープンサイエンスを支えるデータライブラリアン」
<http://www.slideshare.net/ikeuchiuj/2014-41488937>

(図 6)

機関リポジトリによるデータの取り扱い現状

Journal Article(学術雑誌論文)	57289(21.7%)
Thesis or Dissertation(学位論文)	42197(16.4%)
Dissemination Material(発表論文)	27482(10.4%)
Conference Paper(会議発表論文)	22897(8.8%)
Preprint(学術論文前稿)	5633(2.1%)
Book(図書)	19249(7.4%)
Technical Report(学術報告書)	2456(0.9%)
Research Report(研究報告書)	32307(12.3%)
Booklet(小冊子)	46657(17.7%)
Preprint(学術論文前稿)	273(0.1%)
Research Report(研究報告書)	4704(1.8%)
Dataset(データセット)	3343(1.3%)
Software(ソフトウェア)	39(0.0%)
Research Data(研究データ)	17629(6.7%)
合計	263811(100%)

IRDBコンテンツ分析(2015年1月現在):
<http://irdb.nii.ac.jp/analysis/index.php>

(図 8)

合展でそのような話を聞きましたが、実際に合意が得られたところ、やってみようと思ったところから、学術雑誌論文の機関リポジトリ登録に併せた付随データとしての研究データの公開勧誘を始めるという取り組みをしています。

最初のタスクである、なぜ図書館が研究データを扱うのかについて説明します。研究データ保存・公開の意義として、初めに、オープンサイエンスの前提、共有による研究の効率化、研究不正への対応を挙げましたが、そもそもなぜこれを大学図書館がしなければならないのでしょうか。

一般的な文脈では、その下の方に、データリポジトリのニーズがある。つまり、資料の組織化や機関リポジトリ運営の経験を持っているから、あるいは専門家集団による保存の持続性が担保できるからということがいわれます。私は一番の理由として、大学図書館の本来の役割に合致しているということが挙げられると思います。大学設置基準第 38 条には、学術資料は図書館を中心に系統的にそろえられるべきであるという記載があります。それであれば、電子資料であることが図書館が関与しない理由にはほぼならないと考えています。

海外の研究データ管理の動向調査では、機関の研究データ収集のポリシーにどのようなものがあるかを、助成機関から出されているテンプレートを試訳したり、海外の大学でどのように実装したのかを翻訳したりしながら検討を重ねています。

実際に今のリポジトリでどうできるかということで、DataCite や DIF など、研究データのメタデータ標準と Junii2 を比較しています。

システム面では、適切なプラットフォームの検討をしています。イギリスでは DSpace や CKAN といったものがメジャーなようです。これらを使ってみて、本当に使いやすいのかどうか調べています。JAIRO Cloud は今、機関リポジトリの 5 割近くを占めていますが、図書館でやるのだったら、こういうものも使えないかといった検討を進めています。

実際に情報収集をした方がよいということで、2014 年 11 月、図書館総合展にエジンバラ大学のスチュワート・ルイスさんと呼んで発表してもらったり、その前の懇談で情報交換したりしています。2015 年 2 月には、IDCC (International Digital Curation Conference) への参加・発表、サウサンプトン大学、エクセター大学といった研究データを取り扱っている実例のあるところを訪問して、情報交換を行っています。

海外での事例、話などを聞いた限りでの完全な私見ですが、海外での実務はどんな感じなのかというと、ポリシーの策定は、置き換えれば、蔵書構築方針を決めることとほとんど同じです。リポジトリに入れられるデータの種類を決める、収集対象の判断はそもそも誰がするのかというような内容です。

普段の仕事は、メタデータの付与が中心です。電子資料相手の目録業務という言い方もできます。ただ、かなり特徴があるのは、対象となるデータの分野によって、メタデータ項目が大きく変わってくることです。私が所属している極地研究所では、オーロラの解析データか、生物系のデータかによって、入力項目がまるきり違います。一方はメタデータの項目が 155 ぐらいあるけれど、もう一方は 20 ぐらいしかないなど、共通の部分を取れないような代物が多いです。ですから、かなり難しいのです。入力のための仕様書や、マニュアルの整備などは大きな課題になります。また、利用のためのガイダンスは積極的に行っています。

先ほど大学図書館の役割を引き合いに出したのはこのイメージからです。資料の目録と提供は図書館員の仕事の定義に合致します。逆に言うと、今の図書館員ができなければ、似たような組織をつくって、そこに権限を渡すだけという形になります。

あらためてデータ管理が必要だという共通認識が得られたところから、論文の付随データを実際に扱ってみようではないかということで、WG 委員所属大学 (千葉大、岡山大、極地研) のコンテンツ収集活動で検討中です。千葉大では、雑誌論文のリポジトリ搭載依頼時にメールで勧誘を行います。極地研では、極地

研が編集実務を持っているので、そちらで紀要論文に投稿するときに、紀要論文のデータがあるなら出してくださいということを投稿規定に盛り込む、あるいは個別に勧誘することを検討しています。

内閣府のオープンサイエンスに関する検討会で挙げられていた課題は、ポリシー面の検討、システム面の検討、人材育成などです。人材育成というと、具体的には研究者の補助ができる人を育てるということになるため、図書館員だけでいきなりやれというのは難しい話で、具体的にはポストクやURAの方などと協力しながら実務レベルで実現していくのではないかと考えています。

最後に、図書館が果たすべき役割の具体的な検討に入る前に、前提となるアプローチの違いについて触れます。研究者は、最初にデータありきの検討をすることが非常に多いと思います。博物資料や観測・研究データを基に論文や新しいデータを作成するアプローチです。それに対し、図書館員は、整理がメインなので、最初に論文ありきです。論文の根拠としてのデータのみを意識しがちです。リポジトリも今までその流れで来ており、データの一部しか補足できない、補足しづらいので、研究のワークフロー上、使いにくいのではないかと考えています。

結局、自分たちの使いやすいリポジトリを作っているので、リポジトリが林立し、データはどんどん散らばっていく。さらに、そこに相互運用性がないという流れになります。

このモノありきのフローを図書館の実務レベルまで落とすためのヒントとなる先行事例を探すために、私は海外に目を向けていましたが、実は意外と近くにあるということが、このSPARC Japan セミナーの企画WGに参加して見えてきました。現状は、論文は図書館がやる、研究データは研究者のコミュニティがやる、博物資料は博物館と研究者のコミュニティがやるという流れがあります。図書館の外に目を向けると、モノのデータ化やリポジトリの取り扱いの事例は存在します。ですから、ここにヒントがあるのではないかと、研

究フローに沿った運用ができるのではないかと考えています。

もう一つの目標である連携について、あらためて機関リポジトリのクリフォード・リンチさんによる定義を振り返ると、「大学とその構成員が創造したデジタル資料の管理や発信を行うために、大学がそのコミュニティの構成員に提供する一連のサービス」ということで、リポジトリの文脈は「大学」が主語になります。今の状況は相互運用性がないために、リポジトリの林立を招き、どんどん使いにくくなっていて、機関リポジトリの役割を果たしているとは言い難いです。

将来的な目標は、このような定義も踏まえて、研究データにも図書館がきちんとタッチするようにすること、そして、博物資料等も含めた横断的な情報管理体制モデルの構築をこの場で考えることです。そのためには、データに関心のある図書館の人、博物館の人、研究者といったさまざまな方々がこのような場で議論を深めながら、取り組んでいくことが大切だと思っています。

●フロア1 金沢大学の職員です。研究の数値データや、遺跡のデータの公開などに携わっています。研究者というキーワードが何回か出てきましたが、公開されたデータを使って解析するだけの研究者と、人工衛星などで観測して、それを公開する、提供する側にいる研究者は全く別です。今、提供する側のメリットが全くない状態です。データを出した人のメリットが見えるようにしなければ、プラットフォームなどを図書館がやると決めても、結局、掛け声だけで終わるのではないかと考えています。

●南山 まさしくそこは大きな問題だと思っています。データの引用が業績になるなど、データの公開にきちんとしたメリットを与えることをクリアしながらの話だと考えています。

●フロア2 独立行政法人の教員です。

何だかんだ理屈は言ってみるものの、メリットなどしょせんあるはずがないのです。業績にカウントされるといっても、メインは論文でしょう。例えば、このような大きなプロジェクトの中で、データ収集のメンバーでやったというのは、普通、人事の評価をするときにはプライオリティが低いです。中途半端にメリットなどと言ってやる気を出させようとしても、本当にそんなことで人の気持ちが動くという科学的証拠はありませんか。

今まで論文に関して引用数が増える、増えないという議論をしたものの、あまり増えていません。論文でそうなのだから、データはますますそうでしょう。ですから、そのようなメリットで議論するのは格好付けているだけのような気がします。他の議論が欲しいです。

●南山 私の分を超えるので、どなたかこれについてご意見をお持ちの方がいらっしゃったら、ぜひお手を挙げてください。

●フロア3 特にデータ、例えば観測データなどは、公開するまでにキャリブレーションなどで1年かかります。そういうことで、研究者は引用などの業績を求めています。それには科学的根拠がなく、「今から景気が良くなる」と言うのと一緒に、元気を出させるだけになるかもしれませんが、1~2年かかる、論文と違って簡単に公開できないものを参照だけして、みんな論文を書く人に取られていくというのは相当な悩みなのです。その辺は科学的根拠ではありませんが、各業界でいわれていることです。あまり反論になっていませんが。

●フロア2 ですから、業界の判断ではないですか。それを業界横断的な図書館のようなシステムの中で、図書館の人が議論しても、あまり意味がない気がします。例えば実験物理の世界などは、言い方は悪いです

が、非常にうまくやっています。つまり、データを作る人たちが、全然関係ない、例えば計算機のプログラムを書いただけでうんぬんなどと言ってはいけません。傍からはきちんとした報奨システムが作られつつある業界のように見えます。業界のシニアは機関ごとに、雇用についての発言権が非常に強いです。そうすると、そちらの話であって、図書館としては手も足も出ないと考えたらいかがでしょう。

●武田 国立情報学研究所の武田です。データに関しては、分野によって、ものすごく扱いが違うと思います。それは今、ご指摘があったように、論文でもそうでした。まさに著者を誰にするかというような問題です。論文の場合、最低限のライン、例えば調査やサイテーションなどのシステムができたので、何とか共通性があります。しかし、分野を変えると全然違います。ある分野では、研究者が年間1本書いたらすごい人で、ある分野では、100本ぐらい書かないと認められない世界が平気であったりします。論文と同じ名前を付けながら、全く別世界にいるわけです。データの方がもっと多様な状況だと思っています。ですから、分野によって違いがあるというのは当然理解した上で扱わなければいけません。

しかし、一方で、そういうことが業績の一種になってきているというのは、多分我々が決めたというよりは、世界が何となくそういう方向になってしまったのです。ですから、それに対しては、否応なしに研究者、大学はやらざるを得ません。

今の質問は裏を返すと、研究者や大学がやらなければいけないときに、図書館は絡むか絡まないかという質問になるのではないのでしょうか。これは私から図書館に対する質問で、ご意見を聞きたいです。

●南山 私個人の見解として申し上げますと、図書館として業績の話まで絡んでいくかは、当然、図書館員が保存していくものが業績になるということであれば、考えていかなければいけない、少なくとももっておか

なければならないレベルの話であると思います。それに対してどうアプローチしていくかというのは、今、データジャーナルの話などで、既存の論文の評価システムの中に、データを入れ込んで、それを評価基準にしていこうという動きもあるという捉え方をしています。

それをデータジャーナルだから、データジャーナルを今度は図書館で保存する、だからこれは今度、業績としてカウントする。分野的なお話は除いて、そのような流れになっていくのであれば、図書館は積極的に関与すべきであるし、研究データを単体で出した人が評価されるというのであれば、今後、図書館がどのような業務を扱うかによって、関われるか関われないかが変わってくると思います。それは、それこそ図書館が考えることではなく、業界としてのコンセンサスが得られるかどうかに関わってくるのではないかと思います。

●フロア 4 国際農林水産業研究センターの職員です。今の議論は、機関リポジトリで論文を載せるという議論を始めたときも、この間の DRF のメーリングリストなどの議論でも、例えば博士論文を載せるといったとき、物理の分野はこうである、医学の分野はこうであるという、研究者側のコミュニティによって、博士論文のパブリッシュのされ方が違うというところで、割と議論になっていたような記憶があります。

研究データも多分同じで、それが業績になるのかわからないのか、データを作った方がどのぐらいのプライオリティで扱われるかは、コミュニティで違うのではないかと思います。図書館は、このコミュニティではどういう扱われ方をしているのか、業績としてどうなっているのかに踏み込んでいく、つまりコミュニティの中で何が行われているのかをもっと知る必要があるのではないのでしょうか。

具体例を申し上げれば、例えばリンゴの新しい品種を作ったという研究論文の筆頭著者は、最初に系統試験を行った人です。ですから、世に出るまで 30 年ぐ

らいかかりますが、亡くなっても、論文の著者としてきちんと名前が残ります。他の分野ではないかもしれませんが、少なくとも果樹園芸という分野ではそういう扱い方をされるということを知らないと、データを扱う際に、この方が亡くなっていたというとき、その人の評価という話も出てくるでしょうし、何か違う扱い方もあるかもしれません。

それは研究者側のコミュニティでも変わってくる議論だと思うので、恐らく図書館員がしなければならない議論は、サブジェクトライブラリアンまでいかなくても、研究者のコミュニティの中でどう扱われているかということを知りながらデータに踏み込んでいくことです。研究データはそれが分からないと、お手伝いのできない分野ではないかという覚悟を持って今日、聞きにきています。フォローになっているかどうか分かりませんが、そう思っていますがいかがですかという話です。

●フロア 2 今まで言ったことと裏腹ですが、そのように業界の話に取り込まれるのが、本当に図書館員にとっていいのかということをお願いしたいです。確かに、業界ごとに、論文やデータの扱いは違っているし、それがどのような形で業績として評価され、生産した人の人生にどう関わるかについて、いろいろな違いがあることはよく分かります。

しかし一方で、学术界や高等教育界が社会に対して果たすべき責任という議論は十分にあり得るでしょう。その中で特に、こういう学術的な情報の流通が重要な意味を成しているということは、オープンアクセスがハイライトされていることから明らかです。とするならば、そのような学術情報の流通について、これまでも分野を超えておしなべて取り扱ってきた図書館が一定の見識を示して、研究者に対してかくあるべきと述べるのがあってもよいと思いますが、いかがでしょうか。

●南山 研究支援を行うとき、特にある程度特化され

た機関に関しては、その研究者の研究分野に沿った支援の仕方を考えていくべきだと思っています。標準的なものも当然用意しなければなりません、それとは別個で、分野に特化した支援の仕方があってよいと思います。

●フロア2 データをオープン化する意義が語られるとき、多くの場合、研究の公正性・透明性を担保するという理由が挙げられます。公的な資金が入っている以上は、その使用、あるいはその使用の結果が公正であると見せるということが、オープンデータが意味を持つ理由だとされることが往々にしてあると思います。

しかし、業界はほとんどの場合、なれ合いになるので、むしろそれに対する規制的な役割を持ってもいいぐらいではないですか。業界に取り込まれてしまってよろしいのでしょうか。

●三角 それは一定の距離を持った上で、見識を示すという感じですか。

●フロア2 一定の距離は必要ありません。見識さえあればいい。僕は嫌いですが、義務化すればいいだけです。ただ、義務化といっても、もちろんロジックは必要です。「あなたにとってメリットがあるからやりなさい」ではなくて、「そういう立場にいる人が持たなければいけない社会的責任の一部だ」という言い方をせざるを得ないのではないかとということです。

●三角 公的な資金が入っている場合であれば、そのようなロジックが立てやすいと思いますが、そういうことですか。

●フロア2 私的なお金を使っている、国家の安全保障に関わるから見せられない、などの幾つかの例外を除いた場合はそうです。例えば、薬などは微妙な話です。私的な企業が生産するものでありながら、同時に、ポテンシャルには全ての人類に関わる、権利の侵害に

関わるものなので、そのあたりは慎重に考えざるを得ませんが、基本的にそういう線引きはある程度できるのではないかと思います。

●南山 研究支援の話に近くなりますが、図書館がもう少し線引きすべきではないかということを知り、なるほどと思っています。ただ、線引きをするにしても、私が勉強不足などところがあるかもしれませんが、研究の実態を知らな過ぎる、線引きをするには早いという感覚が強いです。線引きをしたいのであれば、研究の全体を分かった上で、しかし、ここまでの社会の要請であるという形で持ってこないか、反発というか、話が全く進まないように感じてしまいます。

●フロア2 それを目指しているということは、要するに実現できないことをやりますと言って、ずるをしていることになりませんか。それよりは、対立を恐れず頑張ってください。

●南山 はい、努力させていただきます。

●三角 他にご意見はありませんか。

●フロア5 高エネルギー加速器研究機構という素粒子の研究所から来ました。私自身、研究者としてデータをたくさん出したことがあります。細かいデータは、「これはこういうものですよ」と図書館の人に1日や2日で説明できるものでもないし、1年かかっても説明できないと思います。それを理解するというのは、どこまでをおっしゃっているのかが理解できません。大量に出るデータをどうしたらいいか、それから、そのデータを生み出すために、例えば私の研究所では、1年間に何十億円もお金を使っています。データをオープンにしてしまった瞬間に、外国の人がそのデータを使って先に論文を書くとか、計算機パワーがたくさんあると、計算がどんどん進んで、先に論文が書けてしまうということもあります。挙げれば挙げるほど

いろいろな問題がありますが、今言ったことについて、どう思いますか。

●南山 一つ目が、研究データをどこまで分かっているのかというお話ですね。これも個人的な話ですが、今まで図書館が図書や雑誌などを整理してきたのも、そもそも中身をどこまで分かっていたのかという議論はいったん置いておいて、情報を流通させるためにこういうものが必要かという観点から物を見てきたと思います。ですから、研究データにおいても、当然コアな部分までは完全に理解できないし、先ほど理解できるのだったらそもそも研究者だというお話も頂いたところですが、それとはまた別個の視点で、情報を流通させるという視点から見て、図書館が貢献できることがあるのではないかと思います。もう一つはどういう点でしたか。

●フロア5 データを公開したら、我々の場合には計算機を使ってデータを解析するので、計算機能力をたくさん持っている集団がいて、自分たちがデータを取ったのに、先にそれを使って論文を書かれてしまう可能性があります。

●南山 そちらに関しては、恐らく国の政策レベルで求められる内容かと思しますので、その辺の動向に関しては後ほど林さんからご説明いただけるものと思います。