

デジタルアーカイブのコンテンツをどうしたら欲しい人に届けられるか（パネルディスカッション）

How can we deliver digital archive contents to the people who need them? (Panel discussion)

2022年9月16日 EAJRS リスボン

Kamiya, Nobutake (Universität Zürich. Asien-Orient-Institut. Japanische Bibliothek)

Egami, Toshinori (International Research Center for Japanese Studies)

Magnussen, Naomi Yabe (University of Oslo library. Humanities and Social Sciences library)

Fukushima, Yukihiro (Keio University)

Iino, Katsunori (National Institute of Informatics)

福島：

（日本のデジタルアーカイブの現状）

日本では博物館法が2022年4月に改正され、デジタルアーカイブについてはじめて書き込まれた。ただし多くの博物館でその準備が整っているわけではない。

公文書については、日本国の公文書は電子が主（紙は従）に変わった。

図書館含め MLA の変化が大きい年となっている。

飯野：

（メタデータの流通について）

「メタデータ流通経路図（学術機関のデジタルアーカイブ）」

(https://iss.ndl.go.jp/information/guideline_attached/ 参照)

各大学・学術機関のデジタルアーカイブのメタデータを集約する仕組みを構築しようとしている。各機関から IRDB に集約して、NDL サーチ・ジャパンサーチに流し込む。機関リポジトリから流し込むルートと、それ以外のスキーマを集約するルートがある。

神谷：

コンテンツにはいろいろな粒度がある。例えば、コンテンツのフルテキストの内容がヒットするとノイズになることもあるのでは。

→飯野：

いまのところ全文ではなく、メタデータが対象。全文対応は今後の課題。

神谷：

オープンアクセスで公開されている書籍のメタデータなどは誰がどこに提供するのだろうか

か。出版者？ 図書館？

→飯野：

電子ブックのメタデータ作成は、図書館では難しい。

電子ブックの総合目録（NACSIS-CATのような）を作る試みを、出版者サイドの協力ですすめている。リッチな書誌データと URL や所蔵館等をデータベースとして提供できるようにしたい。

電子ブックの書誌では、ローマ字を必ず入れてほしいなどの話をしており、海外流通に適したものを考えたい。

→矢部コメント：

今後 MARC システムからの移行に伴いそれに対する対応も必要と考えられる。

矢部：

日本語でしか探せないのは困る。利用者自身の検索でも（日本語熟達者でないと）困るし、また学術図書館においても一般的に日本担当・アジア担当司書の数も減っている。

→福島：

今後ローマ字を極力必須にすべき。

いまの日本のミュージアム業界では、しっかりしたメタデータをつくらなければならないという意識が強すぎたところもあって、メタデータの公開がそもそもすすんでいない。ゆるいメタデータでも良いからまず公開しようという、その意識改革を第一段階にしたい。

→江上：

検索システムに辞書をかませて機械的に対応する方法も（別セッションで歴博後藤さん発言）。

→福島：

人力でなくてよい。あまり厳密さを求めないように。

矢部：

デジタルヒューマニティーズへの貢献のひとつとして、テキストマイニングのためのテキストデータのセット（論文、記事、新聞）を入手しようとするが、日本からの購入価格が高額で、購入が現実的でない。学術機関向けの枠組みなどを、各社からの販売ではなく、国・NDL等のレベルで対応できないだろうか。ノルウェーでは国立図書館に DH ラボがあって、データへのアクセスのサポートをしている。

永崎（人文情報学研究所）：

欧米の日本研究司書の方々からのニーズとしては、簡単に結果が見えるようにデータやツールが提供されてほしいということがあるように思われる（永崎の主観的観測）。データがそろっている分野では研究と成果生産のスピードがはやく、それにつられて学問全体にお

いても若手は速く成果を出すことが求められるようになってきている。日本研究でもツールやデータ・デジタル化資料の充実化が必要なのではないか。日本でもそういうものを整備しなければ。

飯野：

日本の学術情報を海外の人たちにも利用してもらいたい、そのために学術情報・メタデータを流通させたい。ご協力を。

福島：

ミュージアム業界では日本国内に閉じた議論になりがち、外からの議論が必要。

矢部：

自分だけで一方的に考えるよりも、日本や北米の人たちとともに議論して考える機会を持ちたい。

神谷：

ダークアーカイブの構築（購入した電子書籍の蓄積など）をしたい。NDLがDRM付き電子書籍の納本をはじめるとのことなので、活用できるといいと思う。

以下、後藤（チャット）：

飯野氏が提案された統合検索の話につきまして

「全部を検索できるのは本当に良いのか」はちょっと考えるべきなのかなと思います。人間が検索する時には、人間が理解できるような範囲とデータの粒度を作る必要があると思っています。一方で機械で見る時には、機械の粒度を考える必要があるように思います。

メタデータ構築の話につきまして

「マインドセット「から」はそうそう変わらない」という前提でデータ構築を考える必要があると思います。とりわけ日本の博物館は分散型で統一した組織体がないので、「強制力を発生させる」ようなことは極めて難しいです。なので「マインド」とかそういう表現ではなく、機械や仕組みという大きな枠組みを先に作り、そこから流れを考える方向を考えないと、データは増えにくいのではないのでしょうか。

テキストマイニングの話につきまして

このセッションで話題となった、「デジタルアーカイブ」と学術情報流通との間の関係をどう考えるかという視点が必要だと思います。例えば、今回飯野さんが提案されたシステムと、

同じ指向性を持つ情報発見ポータルである Cinii Research との関係はどう考えるか。価格の問題はおそらく電子書籍の価格高騰で考えるべきかという、研究資源共有枠組みになるかと思います。

今回の議論で「プッシュ型の情報提供」や「メタデータそのものではなく、メタデータのメタデータ」こそを見せていくことの重要性を改めて考える機会となりました。すいません、好き勝手に言いまして失礼いたしました。またお話しできますと幸いです。

福島：

>後藤様

マインドセットのところ、そうですね。

ちょっと角度とか戦略とかを改めて考える必要があると存じました。

うっすらと「厳密さよりも早さをほめる文化」とか考えてたましたのでマインドセットと言ってしまうましたが、もうちょっと練ります

飯野：

後藤様 神谷様もおっしゃっていましたが、全文テキストの検索は確かに難しい問題かと思えます。今回のメタデータの収集範囲には、全文テキストは含まれていませんが、いろいろと利用者の感覚に沿った検索結果が出せるように検討を重ねるべきだろうなと思っています。貴重なご意見ありがとうございました。