

知のデジタルアーカイブに関する研究会（第2回）議事要旨

1 日 時 平成23年2月22日（火）10：00～12：00

2 場 所 三田共用会議所大会議室

3 出席者（敬称略）

(1) 構成員

新麗、鈴木卓治（安達文夫構成員代理）、入江伸、植村八潮、大内英範、大場利康、岡本明、小川恵司、植山秀治（加茂竜一構成員代理）、常世田良、杉本重雄、武田英明、鳥越直寿、丸山信人、水谷長志、宮澤彰、盛田宏久、山崎博樹、八日市谷哲生

(2) 総務省・文部科学省

（総務省）原政策統括官、武井大臣官房審議官、安藤情報流通行政局情報流通振興課長、松田情報流通行政局情報流通振興課統括補佐

（文部科学省）高尾文化庁文化財部伝統文化課文化財保護調整室専門職

4 議事概要

(1) 「検討アジェンダ（案）」についての説明。松田情報流通振興課統括補佐より、資料知2-1「検討アジェンダ案」について、説明があった。

- ・ 第1回研究会における構成員からの意見を踏まえて修正したもの。
- ・ 神門構成員からの海外に向けた情報発信、インターナショナルといった観点をという意見を踏まえ、「デジタルアーカイブの構築の必要性」及び「デジタルアーカイブの連携の必要性」について、各デジタルアーカイブが国民、地域社会、さらに国際社会にもたらす効果は何かということで、「国際社会」という文言を追加。
- ・ 新構成員からの意見を踏まえ、4 デジタルアーカイブの構築と連携を促進していく上での課題について、「（4）デジタルアーカイブの構築」に「デジタルアーカイブの長期利用、保守管理面の課題・方策は何か」を追加。
- ・ 4 デジタルアーカイブの構築と連携を促進していく上での課題について、「（5）デジタルアーカイブの相互連携」に「効果的なコンテンツアクセスを助けるためのデジタルアーカイブのデータの横断的な検索・利用手法は何か」、「さまざまなコンテンツとの出会いをより効果的に支援する利用者支援の在り方は何か」と修正。これは、神門構成員から、クエリーを入力する形の検索とは異なる、よりインタラクティブな対話型の検索等、多様な検索・探索のあり方についても検討すべきではないかといったご意見を踏まえ、修正したもの。

(2) 入江構成員より、資料知2-2「(私立)大学図書館におけるデジタルコンテンツの蓄積と提供・課題」に基づき、説明があった。

- ・ 私立大学図書館について、個人的意見も多くあるが、日ごろ業務で感じていることについて説明していきたい。
- ・ 大学図書館の変化について。最近電子学術書というプロジェクトで、出版社の方といろいろお話するが、大学図書館の変化ということについてあまりご理解していただいていることが多いので、少しだけお話しさせていただく。

2000年以降、Googleと電子ジャーナルのインパクトがとても大きい。スライド4にある写真の一番右側にあるように、今、ほとんどの電子ジャーナル、洋雑誌は電子で読めるため、書庫を保存庫に移して、そのエリアを学生用のエリアに開放していく活動は、慶應だけではなく、いろんな大学で今進んでいるところ。

2004年ぐらいからは機関認証をベースにしたリモートアクセスが出ていますので、洋雑誌についてはほとんどOPACから全文が読めるということが当たり前になっている。これをベースにして、今、大学図書館は動いており、空いている書庫を学生用のエリアに開放していこうということが日々起きている。ここの現実的な違いというのがなかなか理解されないが、そういう意味で、今大学図書館は、はっきり言って公立図書館とは全く違うものをご理解いただいたほうがいい。

- ・ 電子ジャーナルアクセス経路について、例えば慶應OPACと言っているものは通常のOPAC。ここから検索して、電子ジャーナルであれば、電子ジャーナルをそこから選び、全文リンクの支援機能があり、そこから慶應統合、いわゆる認証統合に行って、全文を見てくる。PubMedでもGoogle scholarでも同じだが、ウェブサイトでも何でも、基本的にこういう経路の中ですべて動いている。そこには、核に慶應の統合認証機能があって、そこで認証された形で全部が見られる。

Google scholarでも、慶應が持っているものについては、Google scholarを検索すると、@慶應マークが出て、@慶應マークをクリックすると、同じような動きをする。これを当たり前で学生は使っている。

Google等の話でも、アメリカの大学の図書館は既に大量デジタル化の時代に入っている。HathiTrustでは、共同リポジトリを掲げており、これはGoogleのバックヤードで、Googleブックスでデジタル化したものをサーバーに入れて、それを共同利用しようというプロジェクト。それと同時に、紙を集中化し

ていこうということが動いている。2010年ぐらいには680万の本が電子化されている。

- 大学図書館のほぼ30%は電子化が終わったと言われており、毎年10%やると言っているのですが、あと10年かからないで、すべての書籍が電子化される。これをどう利用するかが彼らの最大の問題になっている。
- スライド6ページ目について。現状、どこをデジタル化していくのかというと、電子書籍については、新刊本等は動いていき、パブリックドメインは、国会図書館の方々のプロジェクトで動いていくときに、大学図書館というのは、全体を電子化しないといけないというプロジェクトがあるので、赤い部分のところをどう電子化していつ利用するのか。ここを解決しないと、大学図書館は非常にまずくなるのではないかと考えている。最近自炊の電子書籍が多く、P2Pで音楽がいろいろあったが、これに近い状況は、はっきり言ってすぐ出てくるのではないかと。
- スライド7ページ目は、理工学部の貸し出し統計。左のほうは、統計的に数が少ないが、基本的に理工学部でもロングテール。比較的古い本も使われる。結構古いほうの本がいわゆる貸し出しされていて、特に化学なんかは古いほうが特に使える。この傾向はどこでも同じ。特に法学部も同じような傾向であり、古いほうの電子化がとても大事だということになってくる。
- 電子ジャーナルについては、見てのとおり、英語が圧倒的に多い。日本語はほんとうに少ない。ここをどうするのかということで、慶應の本も、紙の書籍の支出はどんどん減って行って、電子書籍が増えています。ほとんどアメリカではそこは交差しており、50%を超えている。そういった中、現在、京セラ、大日本と一緒に電子書籍の実験をしている。

現在、日本の大学図書館の多くがこの状態になっているということで、電子書籍をどうしていくのかということに入っている。と同時に、一方で、大学図書館は非常に弱体化しており、委託が進み、図書費を減らされ、専門性を否定され、システム能力も減っている。じゃあ、どうするのということまで来ているということが現状であり、コストセンターからはとにかく収益事業をなささいと言われていたが、それはなかなか難しいという現状。

- 基本の取組を簡単に説明させていただく。去年、図書館システムを日本制から海外のものに変えた。これは、メタデータを、電子の資料をどうやってうまく利用するかということ考えてみると、日本のパッケージでは使えないから。そういうものも含め、

今、紙と電子のシステムがある意味分かれていて、紙については、中国国会図書館、延世大学、OCLCなど、いろんなところと協力協定をつくり、電子目録データを使っていたりする。電子データというのは、大体出版社・海外の出版から来るので、それをそのままロードする。そのフォーマットについては、慶應では、和書はJAPAN-MARC、Nacsis、アマゾンから持ってきたりもするが、残りはほとんどMARC21という国際的なフォーマットで運用している。

- 問題なのは、図書館の経費は減っている中、紙と電子の購入費も変わってきているが、運用費について紙のコストが減らない。紙のコストは減らして、電子のコストをどうやってとってくるのかというのが、今、図書館の中での大きな闘い。
- 同時に、MLAという活動もしており、アートセンターと福沢センターとの連携をテーマにしていろいろ活動してきたところ。これはもともと全学の資産を電子化、アーカイブして、共通スキーマをつくらうということから始まったのだが、これはやめた。わかったこととしては、要はメタデータに対する考え方は各組織、コミュニティで異なるということ。それよりもまず相互交換とか、相互利用とか、相互理解を深めようということがまず僕らの最大課題になり、今それをいろいろ進めている。電子化のラインを一緒にするとか、標準をつくるとかということも含めて。
- いろいろ悩んだときに、神崎さんという方が、インターネットはコミュニティの違いを求めるのが本質なので、規制するのはおかしいと随分言われた。それよりも、共通的なものをどう使うかということが大事だろうということで、共通的なスキーマをつくることはやめた。ただし、マッピングはちゃんとやりましょうということで、マッピングをするためには相互理解ということになり、今、相互理解を進めるためにやっているところ。
- 補足だが、図書館の目録担当者はDublin Coreで、メタデータはとれない。なぜかというと、図書館というのは、すごく細かい決まりがあり、誰がとっても同じようにするための決まりが欲しい。Dublin Coreには決まりがないので、これはちょっと難しい。
- もう一つ、周年事業での写真集編纂について紹介させていただく。こういったものが大体アーカイブに残り、相互利用されていくというのが多い。イベントがあり、そこでいろいろなことをやり、それが公開されるということが大きいですが、それを公開するのはとても大変なこと。その際は、事業責任者と著作権の問題が大きい。こういっ

たことが今起きていて中、電子資料とアーカイブのリンクについて、学内で出版している雑誌をいろいろ出しているが、著作権をクリアできていないものは全部黒や白で抜いている。実はこういうのは結構多い。実はここにURLを埋め込みたいと思っている。例えば『窮理図解』という福沢の著作物の説明書について、その絵の中のそれぞれの絵が、福沢の書籍とどうリンクするかということもPDFからURLでリンクしており、そこに福沢の全著作の慶應のサイトがある。

- 最近検索よりも一覧性を上げたいと思っている。Googleから入ってきたとき、そこにまた検索は面倒だと思い、Googleから来たら、それに関する一覧を出したい。それはきれいにしたいというのがあり、例えば福沢が書いた全著作物の表紙を表示して、ここから全文に行くなど。

- 地域との連携については、2000年に山形県と両羽博物図譜というアーカイブをデジタル化したところ。これは慶應の技術がちょうどあり、それを山形でやったわけだが、全文読みだしにして、テキストをつくった。そのテキストをうまく検索に持っていけばよかったのだが、このときは自分の技術がなくて、うまくできなかった。

それについて考えると、例えばGoogleで、松森という両羽博物図譜の著者について検索すると、ウィキペディアに来て、そこから山形のサイト、慶應の研究者などにつながる。そういう意味では、ちゃんと公開すれば、いろんなところがリンクするのは別に難しくはないと思う。これ以上の価値をどうつくるかということ。そこに僕らは何をしたらいいのかということだが、それぞれの団体はそんなに力がなく、お金もないので、単純ではないと思っている。

- 検討アジェンダについて、先ほどの両羽博物図譜で、Googleから検索での露出については、例えば博物図譜で検索すると、酒田両羽博物図譜というのが出てくるが、実は人魚という一番のキラコンテンツが全文に隠れていて引けない。これはやっぱりまずいなと思い、直したいが、山形県には修正するお金がない。アーカイブは維持していくためにはお金が必要だが、イベントでつくったものはそれを維持するお金はつかない。そのうちサーバーがだめになって、お金がなくなってやめてしまう。これが一番大きな問題。

- これをどうやって予算化していくのか、事業を続けていくのかというのが最大の問題であり、これをどうやっていくのかの課題を解決しなければ、つくってなくなっていくという繰り返しになる。

- ローマ字問題というのがある。G o o g l eと一緒にやったときに、向こうから言われるのは、ちゃんとローマ字をつけてほしいと言われる。今さら日本はローマ字なんて要らないのではと思うが、読みがあったときに、ローマ字をどうするかは結構大きな問題であるため、そこは議論してほしいなと思っている。
- ただ、デジタルアーカイブの構築は、必要かと言えば必要だが、先ほどのコストをどうするかという問題が一番大きいということと、無駄なことをどう省いていくのかということが必要。特に連携の必要性は、大体コレクションというのは1個の組織にあるわけではなく分散しているので、これをうまくまとめていく機構は絶対必要。
- 何で進まないかという、例えば大学では、教育・研究が本格的に電子に移ってないからというのが本音であるが、これから教育が電子的手段に移っていくと、編集・利用で重要性が増していくと、そこで使えるコンテンツが必要になってくるので、それをどうやって利用できるようにするかというのが大きな問題になってくる。
- 著作権の問題についてもやはり大きい。例えば学術系の人々の著作物は、大体先生方は公開したい。でも、公開できないことが多く、公開して問題があるかどうかという、多分問題ないが、著作権というものがかぶさっているために公開できないということがとても大きい。
- 運用については、電子というのは、長期保存がほとんど無理なので、組織が維持していく必要があり、マイグレーションを続けていくしかない。紙と違って、置いとくとなくなってしまうものであるから、そういう意味での技術なり組織なりが大事で、そのためには技術の連携も必要。
- 例えば大学に向かって、知のデジタル化、知のために費用をくださいと言っても、絶対来ない。それをどうしていくのかということ、やっぱり今は、例えば1つはコストを省くという意味では、紙での目録コストを省いていくための連携が必要。研究者、出版社と連携をうまくつくりながら、データをつくり流通させる。これは集中するのではなく、きちんと番号を振り、フローを整備して、共同でやっていく。目録などでは、当然出版社は知っているが、読みがわからず調べることになる。こういったものを連携したらいいのではないかとことや、そういった教育講座をやり、使ってもらってお金にするということ、例えば地域でいえば、観光なり、そういうことの中でちゃんと使ってもらおう、お金にするとか、そういうものが絶対必要。
- メタデータフローについて、ライブラリー、パブリッシャー、アグリゲーターなど、

いろいろいて、そこでデータを流通しながらつくっていきこうというモデルがあるが、とにかく電子に行くためには、紙のコストを減らし、組織を維持していくことが大事。ここができなければ、絶対デジタルなんか無理である、そこが一番大きな問題。

- デジタル化について、慶應では人材育成ではいろいろやっており、スライド最終ページをご覧いただきたい。人材育成という意味では、カメラマンを養成したりしているが、貴重書などを撮るカメラマンというのは実は普通のカメラでは撮れない。そういうので規制しないと、貴重書を電子化すると貴重書は壊れる。そういうのを含めて、きちんと対応していかなければならないと考えている。

(3) 八日市谷構成員より、資料知2-3「国立公文書館におけるデジタルアーカイブの取組み」に基づき、説明があった。

- 当館におけるデジタルアーカイブの取組みということで、何をやっているかということを中心にご紹介させていただく。
- 話の本題に入る前に、そもそも国立公文書館について、念のためお話しさせていただく。当館は、日々行政機関で作成される公文書等について、そのうち、歴史的重要性等を勘案されて、長期的に保存していくべきという判断を受けた文書について保管していく施設。昭和46年に総理府の附属機関として設置。諸外国の文書館と比べると、比較的新しいことがわかると思う。その後は平成13年に独法化されて、現在に至っている。
- 独法化される時期に、デジタルアーカイブにかかる事業に着手しており、2つのデジタルアーカイブがある。1つは、アジア歴史資料センターというもので、こちらは、村山政権のときの談話に基づき、公文書館、防衛省防衛研究所図書館、外交史料館が所蔵する戦前のアジア諸国との関係に係る資料をデジタル化していくという事業。もう一つは、国立公文書館デジタルアーカイブで、当館所蔵資料をデジタル化していくという事業。私が、担当しているのは、国立公文書館デジタルアーカイブのほうなので、こちらのお話を中心にさせていただく。
- 国立公文書館デジタルアーカイブのサイトについて、最近リニューアルをしたところ。サービス自体は平成17年から開始しており、現在公文書等のデジタル画像が約868万画像。目録については、基本的に当館所蔵分すべて載っているところ。
- アジア歴史資料センターのほうは、現在2200万画像、デジタル画像が載っており、

いずれも毎年度順次デジタル画像の充実に努めている。

- どうしてこういう事業をやっているかという、勝手にやっているわけではなく、国の政策を受けて推進している。1つは、e-Japanとずっと言われていたが、国のIT政策に書かれている。
- もう一つは、公文書等の適切な管理、保存及び利用に関する懇談会報告というところで、デジタルアーカイブ化の必要性がうたわれ、当館としても対応を進めているというところ。
- 昨年度公文書等の管理に関する法律が成立して、こちらについてもこうした取組を進めていくよううたわれているところ。

こうしたデジタルアーカイブ化について取り組んでいるところだが、「デジタルアーカイブ化推進に関する基本的な考え方」の中に、例えば①コンテンツの流通、発信ということがある。国のIT政策でも、情報発信力の強化とずっと言われている。また、一般の皆様、国内外を問わず、館所蔵資料を利用できる環境を整備していくことが一般的に必要と思われている。また地方公文書館などの関係機関とも随時連携していきましょうということも考えている。

- 最近顕在化してきた問題で、現在私も対応しているところだが、電子的な公文書の保存、利用に向けた取り組みということで、現在鋭意取り組んでいるところで、そうしたものも、従来のデジタルアーカイブという枠組みに加えて、電子公文書への対応として入ってきている。今申し上げたのは②の点。③将来像ということについて、こういった研究会で議論されるとおり、いろいろな連携の中から知を生産していくということもあると思うが、当館としても、デジタルアーカイブ推進要綱というものを検討したときに、そういった将来性というものはあるんだなということ認識して動いている。
- 当館のデジタルアーカイブのポイントについて。基本的には、「いつでも、どこでも、だれでも、自由に、無料で」使っていただくというコンセプト。利用と保存という、我々、そうした目的を持つ機関でございますので、そういったコンセプトに合う仕組みを入れていきたいと思います。また、データについては、もう今さら言うまでもないが、世の中、情報化社会であるから、そうしたところで利用いただける情報を提供していくということを考えている。
- そのために採用している技術的な仕様とかフォーマットについて、(2)特徴とい

うところで書いてある。

- 国立公文書館デジタルアーカイブのトップページから、当館が所蔵する文書を検索してもらい、検索結果から文書のデジタル画像を閲覧いただける仕組みになっている。それは基本的な機能として必要なものだと考えているが、それだけではなかなか一般の方には使い勝手がよろしくないんじゃないかということもあり、マウス一つでブラウジングしながら画像を探していただける場所も画面の下に設けており、ここで重要文化財であるとか、貴重資料であるとか、すぐに見られるようになっている。

こうしたデジタルアーカイブについて目録データを作成して、検索いただけるようにしているわけだが、アーカイブの目録で特徴的なところとしては、例えば今画面に当館の主な資料になる公文書と内閣文庫の絵をつけているが、公文書の場合は、必ずどこでだれが何のために作成した行政文書なのかという情報を持っているので、我々から見ると、移管省庁になるが、移管省庁の何々局の何々部の何々課が作成した文書というものを資料群構造として階層化して保存している。なので、メタデータ、資料群データのつくり方についても、そういった組織情報が階層的に構造化されている。内閣文庫については、基本的にこれは出版物が多いので、深い構造というのはあまり持っていない。

- いずれにしても、現在EADという公文書館の世界でよく使われている技術上の標準のメタデータセットに合わせて提供しているので、いろいろな機関で利用いただける仕組みを整えているところ。
- そうした仕組みに乗って、どういう画像が流れているかということについて、あまりそう一般の方はすぐ見たいと思わないかもしれないが、行政文書がある。また、明治期の錦の御旗があったり、日本国憲法の画像があったり、江戸時代に作成された国絵図があったりと、多種多様な資料がデジタル化して、掲載されている。
- 特に当館に所蔵している所蔵資料、今、120万冊ぐらい資料数あるが、一遍にすべてをデジタル化するのはなかなか難しいというのは当たり前であるが、そういうこともあり、できるだけ国民の皆さんに使っていただけるような資料や、劣化が激しい資料であるとか、物理的に使っていただくのが難しい重要文化財資料であるとか、そういった資料を中心に順次デジタル化しているところ。
- そうしたデータについてはいろんなところで使っていただけるようになっていると先ほど申し上げたところだが、横断検索の仕組みであるとか入れているので、現在

こちらに書いてあるのはアジア歴史資料センターであるとか、N I I 国会図書館、岡山県立記録資料館、奈良県立図書情報館などと今データのやりとり等できるようになっている。

- これが当館で現在進めているデジタルアーカイブの一番基本的なもの。それだけではなかなか普及していく分にはちょっとつらいということもあるので、ウェブサイトで、デジタルアーカイブで使っている画像などを用いながら、月がわりのコンテンツをつくったり、季節ごとのデジタル展示をつくったりと、いろいろ工夫をしており、ぜひ皆さんに使っていただきたいと思っている。
- 今回の研究会が連携というところに主眼を置いているので、当館でもこうした取り組みをしておりますということをご紹介させていただく。

平成19年にさかのぼるのが、当館そのものがデジタルアーカイブ化を進めるとともに、地方自治体の公文書館においてもデジタルアーカイブが何とか進まないものかと考え、まずは調査を始めながら、じゃあ、どういった取り組みができるかと考えて、いろいろ検討した結果、現在、題目にあるように、デジタルアーカイブの標準仕様書という文書を作成して、各館に、時々私自身回って歩いて説明をしまっているところ。

これまでの取り組み、簡単に画面に出しておりますが、いろいろアンケート調査をやったり、訪問調査をやったり、取り組んできたところ。いろいろ聞いてみると、どこも人も予算も何もないという実情もなかなか厳しいものがあり、その中で、当館としてできることといろいろ考えると、技術的な支援をやっていくことがある。それで、デジタルアーカイブをつくるためのノウハウを集めた文書をつくり、それをまずは提供してみたらどうかと考えて、取り組んできた。

- 文書については、大分前に作成しており、システムを実際に組んで、マニュアルも作成し、平成21年から各自治体と基本的には公文書館に出向いて説明してきている。
- 当時どういう調査をして、どういう答えがあったかという、基本的にどこの公文書館でも、マイクロフィルム化とかデジタル画像化というのはどんどんやっていきしたが、進んでいないという回答であった。

どこでもホームページで所蔵している資料を住民のために紹介したり、そういう取り組みはぜひやっていきたいと考えているということだが、いろいろな課題があって、目録情報の提供、画像情報の提供は非常に難しいということであった。

そういったこともあって、現在公文書館といっても、全国に2000近い自治体があったと思うが、公文書館があるのは50ちょっと。それらの自治体が持っている文書の情報については、今ばらばらというのが現実ではないかと思う。こういった状態を何とかデジタルアーカイブの仕組みを入れていくことで、情報を非常に探しやすい、いろんな県の文書が他県でも探せる、住民がすぐ自分の県の文書を探せるようになる、こういった取り組みができるんじゃないかと考え、取り組んでいる。

- ・ 標準仕様書については、例えばシステムをつくるときはこういう基本的な考え方を使ってはどうかということを入れており、ホームページに載せて、誰もがご自由に利用できるようになっている。こうした取り組みがすぐに結果が出るというものではないと思うが、地道に取り組んでまいりたい。

これについて、以下のやりとりがあった。

- ・ いわゆる書かれた文書のテキスト化、全文データについてはどのように取りくんでいるのか。(武田構成員)
- ・ 基本的には文書の画像化をしており、テキスト化は、話はよくあるが、事実上手が届かないというところ。(八日市谷構成員)
- ・ 公立公文書館で持たれているものは、各府省から移管されてきた文書類で、例えば先ほど出ていた吾妻鏡など、ほんとうに古いものや江戸幕府からの移管文書などもある。一方、最近の公文書管理法の関係でいうと、情報公開法も関連するが、役割の違いとして、各府省が文書管理して、公開をする際のある種アクセスの仕方、メタデータ管理の話と、それと実際に公文書館に移管されてきてからの文書の管理の仕方というのは、そこで違いが当然出てくる。その間でのつながり、全く今は切れているのか、あるいはある程度、もともとの府省でつけられていた情報、メタ情報、それを参考にされているのかという点と、国立公文書館以外の都道府県での公文書館だと、そのあたりのつながりはどういうふうになっているのか教えていただきたい。(杉本座長)
- ・ 基本的な一般論として答えると、アーカイブの世界では、お預かりする資料に書いてあるメタデータ、もともとある情報をそのままお預かりするというのが一般論としてある。そういった観点から考えれば、行政文書の場合も、基本はそこに書いてある行政文書ファイル管理に置きかえてあるものをそのままきちっと入れていくというのがまずは基本としてある。その先の利用性についていろいろ考えていくことは

また違う話だと思うが、まずは基本的にもとある情報をきちんと残すというところがあるのではないか。(八日市谷構成員)

- 国の組織は非常に大きいので、全部統一するのは現実的には難しいと思うが、いかに横並びで見られるかということは、よく求められる要求である。(杉本座長)
- 公文書に関して著作権に関してはどのような考え方あるのか。現状の扱いは。(武田構成員)
- 検討されているところ。今ここで出したものは、基本的に行政文書がメインであり、さらに、かなり昔のものばかりなので、そのあたり、抵触しないのではないかと考えている。(八日市谷構成員)
- 今、デジタル化して公開している部分に関して、公開のときの著作権の条件は何かつけているのか。公開されるものは、一応ある種のパブリックドメインという扱いになるのか。(武田構成員) 公開
- パブリックドメインという扱い。(八日市谷構成員)
- 基本的に国がつくったものなので、それに関連する資料で国がつくったもの以外のものというのが含まれている可能性があるが、それに関しても、特例的なことになり得るということを知っている。(杉本座長)

(4) 大内構成員より、資料知2-4「史料のデジタルアーカイブとその課題」に基づき、説明があった。

- まず史料編纂所というのがどういったことをしているのかということの説明して、現在取り組んでいるデジタル化について、それから、対象が史料であるということの持っている問題点というものについて少し申し述べたいと思う。
- 史料編纂所は、「国の内外に残されている古代から明治維新に至る日本の歴史に係る史料を蒐集、研究するとともに、その成果をふまえて、日本史研究の基礎となる史料集を編纂・出版する研究所」という位置づけ。史料編纂所のルーツは、HP上での公式見解としては、江戸幕府によって援助を受けて、塙保己一が開設した和学講談所にさかのぼるというもの。それを信ずるならば、200年以上の歴史を持っているということになるが、1869年、明治に入ってから、史料編輯国史校正局というものがつくられ、そこで近代的な歴史記述を始め、そこが直接的には本所のルーツと言っていると思う。

まずは、ペリー来航以来の諸藩あるいは華族が持っていた日記、手控えのたぐいを提出させ、それらをまとめるところから始めたそうであり、明治国家の建設と密接につながっていたのであろうということが想像される。

その後、**1888年**に帝国大学文学部の一機関となり、その後**1950年**に附置研究所となり、文学部から独立して今に至るといのが簡単な史料編纂所の歴史。

- ・ 史料集の編纂・出版というものを目的としているので、史料蒐集ということがまず第一。国宝の島津家文書、国の重文である実隆公記等の原本史料をはじめとして、そのほか、国内外で所員が出向いていき、個人のお宅のこともあり、さまざまな機関のこともあるが、史料の複製をつくらせてもらう。現在ではマイクロフィルムでの撮影、あるいはデジタルカメラでの撮影ということをさせていただき、複製をつくらせていただく。それを所蔵していくというのが現在の大きな事業の柱。
- ・ マイクロフィルムが一般的になる前には、そこにあるような影写、謄写、模写と書いてあるが、要するに書き写し。それによって各地の史料を、実際に所員が行き、あるいは史料をお預かりして、実際に書き写すというようなことで史料を集めていた。

近年では書き写しの行為自体が研究対象としておもしろいということにもなっているわけだが、膨大な複製史料を作成して、それを架蔵しているということ。

- ・ その複製史料をもとに史料集を刊行するというのが、史料編纂所の目的。その史料については、本所の図書室、あるいはウェブサイトにおいて公開されている。ウェブサイトのほうは、他機関、個人の所蔵になるものについては、制約はあるが、本所に来ていただければ、原則として大体のものはごらんいただける。

また、**2006年**から前近代日本史情報国際センターというものが本所内に設置されており、情報技術を用いた史料の公開というものに取り組んできたわけだが、**2009年6月**研究拠点に認定され、**2010年度**からはより一層、本所がこれまで作成してきた複製史料をはじめとする所蔵物の公開に力を入れていくこととなった。

- ・ 現在の取り組みの一例を申し述べたいと思うが、我々、出向いていき、写真を撮らせていただくのを採訪と言っている。現在では最初からデジタルカメラで撮ることもあるが、まだマイクロで撮っている場合もあり、採訪マイクロと我々が言っているものがどんどん所内にたまってくる。ただ、これをデジタル化しないといけないということが必然的な流れになっている。それはまずはマイクロフィルムの劣化という問題。これにつきましては、最近のマイクロフィルム、PETと言っているものについては、

ちゃんと管理すれば500年はもつと言われているので、そんなに心配ないのかもしれないが、ちょっと前にあったTACと言っていたマイクロフィルムについては、史料編纂所の書庫の中が酢酸のにおいが充満していたと。湿気と反応して酢酸が発生し、マイクロフィルムが傷んでしまったというようなことがあったらしくて、それをよりいい質のマイクロフィルムに置きかえるということもやっているが、本所では基本的にマイクロフィルムは、印刷をして、写真帳というものをつくる。それが多いたときには年200リールほど、1リール600コマとすると、12万コマということになると思うが、大体毎年1000冊ぐらいの写真帳をつくり、架蔵していくということをしてきた。していたというのは、CH印画紙というものが、2008年だったと思うが、生産を中止すると業者から通告があり、多少残っている在庫で当面はしのいだが、それができなくなってしまった。つまり、地方に行って撮らせていただいたマイクロフィルムを持ち帰ってきて、写真帳をつくるという一連のフローが成り立たなくなってしまったというようなこと。

- ・ 採訪マイクロが、どれぐらい本所にはあるかということ、大体500万コマぐらいじゃないかと。ぐらいじゃないかというのは、全貌がいまだによくわからないところが実はあり、でも、大体そのぐらいだろう。

うち80万コマ前後をとにかくデジタル化しようじゃないかということになり、大きい科研をとって、これはデジタル化だけが目的の科研ではないが、その科研の中でデジタル化を進めていこうということを取り組んでいる。今年度までで約200万コマ、デジタル化した。残り2年あるが、全部デジタル化できるかどうか微妙だということころ。デジタル化して、1コマ1コマにメタデータをつけて、簡単なものではありませんけれども、検索システムをつくらうということを進めている。

- ・ デジタル化における問題点、幾つか直面しているところを申し述べるが、まず歴史史料にメタデータをつけるときの細かさ。これは本所の所蔵史料目録におけるメタデータの数は、だれがいつ入力したかといった管理項目も含むが、書目54、冊36、細目27と、文書1点に対して100程度のメタデータ、項目を付与していくという作業をやらなくてはいけない。ここでいう書目、冊、細目というのは、細目というのが文書1点だとしますと、冊というのはそれが束になっている状態。そして書目というのが、例えば島津家文書といったような総体をあらわすもので、全体としてかなりきめ細かいメタデータを付与しているということ。これは既存の規格に合わせているというこ

とではなくて、文書の多様性、多様なあり方に対応してつけているということ。

とはいえ、他機関との横断検索等の連携というのも視野には入れているので、このメタデータを整理して、既存の規格に合わせていくような研究もあわせて行っている。

- そしてもう一つ問題点として、マイクロを画像化しているわけだが、どの程度の画像で持つかというのが大きな問題。画像をつくれればつくるほど、保存する場所が必要になる。現在史料編纂所には、大きな科研のものも含めると、50テラバイト以上のストレージがあると思うが、これを毎年毎年増やしていかないと、画像化に追いついていけない状態で、この経費も大変なことになっている。とりあえずJPEGファイル、これは広く一般的なフォーマットであって、400PPIというのは、いろんなご意見あると思うが、一応編纂というのが一番の目的ですので、編纂に必要な程度の精彩さというもので、この程度の画像をつくっている。
- デジタル化をした後、公開するときの問題点、幾つかあると思うが、本所所蔵史料については、あまり制約はない。問題は、先ほど申し上げた採訪によって得られた複製、所外史料とでもいうのでしょうか、そういったものを公開する場合の問題点で、採訪は数十年前からやっているもので、例えばウェブページを通じて公開をしますよということを実際にお伝えしないで、マイクロフィルムにとってきているもののほうが多い。それを1点1点採訪先にウェブページから見られるようにしてもよろしいでしょうかというようなお伺いを立てていく作業が必要だろうと思っているが、これが非常に大変だということ。史料の所在が変わっていることも多く、採訪先がなくなっているということもある。
- 先週も、20数年前に所員が採訪に行き撮らせていただいた写真が本所にあるが、県の方が史料編纂所にはこういう写真があるけれども、現在はその所蔵先がもうないと。個人のお宅だったが、その家自体がなくなっていると。どこに行ったか知らないかということで、追跡調査をされている方がお見えになったが、本所ではわからなかった。そういった困難な作業もあるし、また採訪先によって要望が微妙に異なることも多く、住所の公開をしたくないというのは多いが、閲覧室での閲覧はいいけれども、ウェブ公開はしないでほしい、あるいは、この史料を持っていることそのものを知られたくないというような方も中にはいる。
- そこで、本所では、現在ウェブ閲覧、あるいは閲覧したものをダウンロードできるかどうか、また印刷できるかどうか、そういった段階別に権限の制御をするようなコ

コンテンツ保護プログラムの導入を進めておりまして、所員の中でも複数のカテゴリーがあり、このカテゴリーの所員はわりと制約なく何でもできる、このカテゴリーの所員は印刷はできないとか、いろんな形でコンテンツ保護をしている。

- 本所所蔵史料及び所外史料のデジタル化と公開、その問題点ということについて、一般的な史料館、文書館さんにおいては、今八日市谷構成員からもあったように、特に地方でのこうしたデジタルアーカイブの事例というのは、私の見聞きする範囲だが、少ないのではないかと。結局はカネとヒトの問題なのかなということになるわけだが、そこで史料編纂所では、史料編纂所のデータベースを使っていただくということも可能ですよというようなこととお話しさせていただくこともある。

例えば県史あるいは、市史の編纂によってかなりたくさん史料が蒐集されるが、編纂後、広くそれを公開しなければいけないといったときに、どうやったらいいのかというようなところで、史料編纂所の所員が大体は県史とか市史の編纂には加わっていることが多いので、相談を受けることがあるわけだが、史料編纂所のデータベースにとりあえず蒐集された史料のメタデータだけでも登録したらどうですかというようなことを申し上げたり、あるいは、地方の文書館さんがお持ちの史料、それを編纂所のほうで写真を撮らせていただき、それをデジタル化した場合に、その文書館さんに、史料編纂所のものでいいが、PCを置いて、そのPCから史料編纂所の画像サーバーにアクセスをして、これがうちの持っている史料ですよというのを、その文書館さんのほうでござんいただくというようなことも可能。そういったことも提案させていただくことも今やっている。

- 対象が史料であることによる問題点というのを二、三申し述べさせていただく。例えばどこかの村の庄屋さんか何かは大量の文書を残されていたと。それを編纂所のほうで調べさせていただく。そういうことはよくあるが、大体は文書の登場人物というのは、現在の持主のご先祖様になるわけですが、調べていくと、ご先祖様のこういうところは公開されると困るなというような文書が出てきたりすることが中にはあり、所蔵者の方が公開を拒まれるというようなこともある。

また、特に近世以降の資料になると、差別に関する内容を含むような資料の中にはあるし、中世以前だとそれほど問題ないのですけれども、例えば絵図などをとりましても、そういったものに関係する地名が書かれている絵図は少なくない。そういったものを公開させていただいていいのかどうかというようなことに関しては、編纂所

内の研究者個々の立場、編纂所としての立場をどうするのか。これはまだ所内で意思統一されていないというようなどころもある。

対象物が史料であるというところに起因するそういった問題も幾つか存在するということ。

これについて、以下のやりとりがあった。

- ・ 国立公文書館のデジタルアーカイブの連携という可能はあるのか。(入江構成員)
- ・ 技術的にはもちろん入ることは可能だと思うが、内容的にどうか。関係があると見る人もいるだろうし、これはちょっと関係ないんじゃないかと思う人もいるだろうと思う。ただ、別に関係ない者同士が連携して別に悪いことはないと思うので、そういった関係ない者同士で連携していけば、最終的には1つの大きな共同体になると思うので、興味はもちろんある。また、共通するところもおそらくあるんだろうと思う。(大内構成員)
- ・ 現状ではそういうこととは全く無関係だということか。(入江構成員)
- ・ 公文書館とは、私の知る限りではあまり連携のお話はしてないんじゃないか。(大内構成員)
- ・ 大内構成員からお話があったととおりで、広い意味で関係するものだと考えているし、今後そういうお話があれば可能だとは思う。(八日市谷構成員)
- ・ 絵図などの公開ではばかれるものというのは、それに関してどういう議論されているか、興味が少しある。それから、公立公文書館さんも同じようなことを多分議論されて、幾らかそういうものもあえて公開されているような部分があると聞いているが、その考え方を参考のために教えてほしい。(山崎構成員)
- ・ まず申し上げるのは、所としての方針はまだ形成されるに至ってないということ。ただ、所員個人個人が歴史学の研究者であるから、個人の立場としては所員それぞれがいろんな考えを持っている。どちらかという、私の知る範囲では、歴史的な事実として公開に前向きな意見を持つ所員が多いのではないかなと思うが、所としてはまだ慎重に対処している。(大内構成員)
- ・ 資料の中身までは担当していないので、正確なお答えはできかねるが、そういった資料がある可能性は否定はできないのはあると思う。事実関係として、今、それが出ているかどうかというのは確認しなければ何とも申し上げられないが。別の部

署で議論している可能性はある。(八日市谷構成員)

- ・ 個人情報だとはっきり決まっているが、そこは難しいところ。それと、史料の種類にもよる。

(5) 意見交換

- ・ 大内構成員への質問であるが、メタデータを考えるときに、いわゆる I S A D、E A D などの標準等の関係については、検討するのか。(入江委員)
- ・ 検討する。大分前からその辺については内部的には研究を進めている。本所のメタデータがあまりにも細か過ぎるので、合うところだけ合わせていくような形で、90年代の後半ぐらいから研究している。具体的には、まだ実現には至っていないが、人間文化研究機構の研究資源共有化システムとの連携を前提に進めさせていただいているところ。

(大内構成員)

- ・ この検討アジェンダの中では、制度、資金面の課題に関係するが、今の大内構成員の話にもあったように、公的機関がデジタル化することと、少なくとも著作権ではないわけだが、所有者の権利というのは厳として社会的にはあり、公開不可と言われることもある。これに関してどういうサービスするならこれは許すというようなことを少しパターン化して、権利を持っている人々に対してどこまでやっていいですかと聞きやすくするような整理を行うことを考えてもいいのではないか。(宮澤委員)

技術的にデジタル化されたものをデジタルでコピーできないようにする公開、所内だけの公開、印刷、無制限にコピー可能など、そういう段階をパターン化するというのを考えてもいいんじゃないかと思う。

- ・ 美術品について、著作権が切れているものがあつたとした場合、それをカメラで撮って公開するというのは、著作権上問題ないんじゃないかと思うが、それは問題あるのか。

(入江委員)

- ・ 考え方はいろいろあると思うが、国立美術館の場合、所蔵品を持っているのは4館。ほとんどが著作権があるので、それはともかくとして、古美術については、所有者の許諾をとる必要があるかどうかの議論については、ほとんど大昔、30年前、大和文華館の事例がある。大和文華館は、所有について、利用の制限をフリーにしたという事実や、「顔真卿自書建中告身帖事件」裁判の判例でもあつたと思う。もし所蔵にかかわることで、そのときはウェブの話はなかったが、論争になると、その判例を持ってくれば、持

っていることについての所蔵権について主張は難しいと思う。しかしながら、慣例として、博物館、あるいは関係機関は所蔵者に対して非常に手厚い配慮を行うので、それは一例を踏まえて、コンテンツを利用していくということになると思う。

先ほど著作権、そういうことの処理についてのパターンを整理したほうがよいというご意見があったが、それは全くそうだが、例えばデジタルアーカイブ推進協議会は十数年活動してきて、2006年に解散した。推進協議会においては、著作権処理についての、どういうふう処理をしていけばよいかというロードマップも出している。なので、この会議全体が、およそ20年ぐらいデジタルアーカイブについてめぐった議論を、またゼロからやるのかという感じはある。デジタルアーカイブ白書というのが少なくとも5冊出ているので、その成果を踏まえた上で議論をしていくのも1つの方策ではないか。

いろいろな議論はあると思うが、1つ、美術館のサイドから、簡単にお話しすると、この20年ぐらい、2つの大きなデジタルアーカイブ構築についての山があった。1つは、ハイビジョン。ハイビジョンのとき、非常に盛り上がったが、皆さんご存じのように、いろんな理由があって、あれは頓挫した。

それからもう一つは、デジタルアーカイブ推進協議会が組織された段階、それは90年後半から00年代だと思うが、それから2006年までの間、デジタルアーカイブについての議論は多方面において行われた。

そうすると、2006年から現状の2011年のこの5、6年の間に何が進展したのか、どういう技術的な要素、あるいは社会的な要素においてデジタルアーカイブ構築がしやすくなったのかということ整理をする。それから逆に、デジタルアーカイブ推進協議会がなぜ解散したかも含めて、この5、6年、逆にデジタルアーカイブをつくろうという盛り上がりは、ミュージアムにおいて、特に美術館においてはダウンしている。もろもろの要因もあるが。

そこを整理した上で、技術的、社会的にデジタルアーカイブを構築するについてのアドバン스는何なのか、逆に後退した要素は何なのかということ整理した上で、でも、全体からするならば、デジタルアーカイブをつくる環境は、昔に比べれば敷居はかなり落ちているはず。落ちているにもかかわらず、なぜ推進できないのかということ次回ミュージアムの立場からお話をしようと思うが、おそらくきょうは図書館、それからアーカイブのお話が出たが、それに比べれば、ミュージアムのほうがデジタルアーカイブを構築することについての意義と効果は、世の中一般に対して非常に見えやすいと思う。

しかしながら、なかなか構築ができていない。その理由は、きょう八日市谷構成員からお話しがあったが、やはり人の問題、それから予算的な問題というのが、あるプラスを補てんすれば、昔よりはずっと構築の可能性があるのに進まないということと、それからもう一つは、今回の知のデジタルアーカイブに関する研究会の最終的な、あるゴールを設定する。あるいは、横断的な検索を可能にするとするならば、どこがプラットフォームになるのかというのをある程度明示して、それが例えば国立国会図書館のPORTAかもしれないし、あるいはどこかかもしれないが、そのイメージを明確に出さないと、デジタルアーカイブ推進協議会が2006年に終わったと同じようなことになりかねないので、そういうことを整理して臨まないと、この会議はまた10年前と一緒にではないか。

(水谷委員)

- ・ 以前私も、デジタルアーカイブをつくるとき、ちょうど今のお話と同じようなことで、必要性をなかなか理解していただけない。つまり、アーカイブをつくるということは、そこに来なくても見られる状態につながっていくということ。民間団体などに貴重な史料があり、それをデジタル化したいというお話を持っていったときに、それはかなり言われた。デジタル化することによるメリットがわからない。つまり、そもそもあれば来ていただけるから、そのほうがいいのではないかと。そこに我々、価値を持っているというお話をされる。マイグレーションのみ行って、現在でも実際公開までは至っていない。その当時は10年前なので、そういう状況の中で、そういうことを理解していただくというのは難しいと思っていたが、現在大分状況が変わってきて、前回私がお話ししたように、公開することによってメリットが明らかに出てきている状態なので、必要性というものをしっかり明記して、そのことによってむしろ史料の価値が上がっていくんだということを考えていかなければ、一部には、来てみるべき、あるいはウェブ公開は嫌、閲覧のみにしようとかというのは、先ほどのお話の中にもあったが、やはり同じようなことがあちらこちらで起きている。美術館もそういうところがあるかもしれない。なので、そういうところの有用性というものをかなり強調してこの会議の中で示していかなければ、技術的な問題はもちろんあると思うが、なかなかそれだけでは解決できない問題がそこにあると思う。それは多分、専門職員の確保の問題だかと、そういうことにもつながっていく。(山崎委員)
- ・ この研究会が設置されている懇談会が、もともと電子書籍に関することを扱っている。なぜ、そういうことが出てきたかということ、この10年ぐらいの間における情報環境の大

きな変化であると思う。新しい情報環境の中で、例えば美術館、博物館というのは、ある種出版コンテンツのさらにそのもとを持っている、そういう組織かと思うので、そういう一番の上流の立場として、非常に大きなコントリビューションがなされることになるかと理解しているし、それから、逆に図書館というのは、ある種一番の下流にいるのかもしれず、利用者に最も近いところに、コンテンツをある種集めて、アクセスポイントとして、何かを探しやすくするという、そういう立場が図書館にはかなり求められる。そういう意味で、MLAで、今、頭文字を並べて言っているが、もともとかなり役割が違う。それは入江構成員の話でも、コミュニティーの違いがありますよねという、そういうお話もあったところにも反映しているのかなと理解して、自分自身は聞いてはいた。基本的には情報環境の変化、それから、特に携帯型のいろんな端末が出てきて、どこにいてもネットにつながっていて、どこにいてもコンテンツにアクセスできる環境ができてきた。そこで新しくデジタルアーカイブなるものがどんなふうに使えるようになっていくべきか、どんなふう在世の中に役に立っていくべきかと思う。

そういう意味では、前回のお話でも、いわゆるコマーシャルなコンテンツとの境目というのはどうかという議論もあったかと思うが、権利管理の関係の話も出てきて、ここでは、権利管理そのものは扱わないけれども、いわゆるコマーシャルな出版物というのを除外して考えることはしないようにしましょうということで、ここでは合意したと思う。

なので、新しい情報環境においてデジタルなコンテンツ、あるいはデジタルになっていないコンテンツにしても、できるだけうまく効率よく利用者に届けるにはどうしているのかというのを考えればいいのかと私自身は思っている。そういう意味では、メタデータに関する議論というのは、ここで、僕自身のほうが関心のあるところではあるが、とても大事な役割を持っているかなと思う。(杉本座長)

- 要は90年代の電子図書館というのは、予算がおりてきて、電子化して、ウェブに出してというモデルはもう終わっている。どうやってそういう仕事を続けていくかということになっていて、人とコストの問題とプラットフォームの問題は避けては通れないと思うが、それを言っても結局何も進まないときに、杉本座長がおっしゃったように、どうやってそのモデルをつくれるかということだと思っていて、大学図書館でいうと、教育コンテンツ。教育コンテンツをつくるときには、どうしても教育が電子に行くと、電子的なリソースなりが必要になってきて、その編集をしないといけない。編集するときに、

例えばミュージアムの資料、文書館の資料、図書館の資料などいろんな資料を使うときに、それを使いやすくしないといけない。そこのビジネスなりのコストの中で、何らかの電子化なり、知をつくっていくことをしないと、無理なのではないか。知という言葉だけで法人に説明しても、そんなもの、絶対金はでてこず、教育コンテンツとしてこれが必要ですよということや、ここで金もうけましようという言葉を使わないと、結局図書館的には仕事ができないという状態の中で、コストとして今あるいろんな管理コストを減らしていきながら、私立大学図書館としては、多分将来できなくなってくるだろうと思っており、その突破口みたいなものをどこかであげないと、次はないのかなと思う。先ほど大内構成員がおっしゃったように、僕らも今50テラのサーバーを持って、毎年2テラ以上ずつ増えていく。それは図書館の仕事ではできない仕事なので、そういうコストもどうするか。クラウドといっても、リスクは高いので、そういうことをどうしていくかというのは、国会図書館がサーバーくれないかなど思う。(入江委員)

- ・ 商業的なものに関する議論も排除しないというお話を伺ったので、今ここで考えたほうがいいかなと思うことに、入江さんのお話に、大学図書館はここまで来ている、特に電子ジャーナルについて、検索から現物のダウンロード、印刷まで、普通にちゃんといけるんだと。その点で公共図書館とは全然違うというお話をされたが、そうなってきたときに、大学に所属しない人がそのようなものにアクセス、かえってできなくなっているのではないかという問題を感じる。

以前ならば、紹介状なり何らかの形で大学に行ってアクセスすればよかったわけだが、今公共図書館で商業的、学術的な電子ジャーナルを導入しようとする際、市民全員にサービスするといったら、高くて買える値段ではなくなる。そうすると、そういうものに対して、公共図書館、あるいは市民一般がアクセスする道がかえって狭くなっているのではないか。(宮澤委員)

- ・ アメリカなどの例を見ると、多くの場合、例えばカウンティ単位でコンソーシアムをつくって、大学と専門と公共と学校図書館がコンソーシアムをつくり、そこでコストを一元化して、ディストリビューターやベンダーとコンソーシアムが値段の交渉をする。商用のデータベース、電子ジャーナルを200も300も公共や学校図書館も含めて同時に使っている例なんていうのは非常に一般的。そのときに、アクセス数で値段を設定したりするので、大学のアカデミックプライスと比べると、公共の場合は非常に安いコストで実際は一般の市民は無料で使えるというのは、非常に当たり前の状況。例えばブルー

ムバーグの端末は、日本では証券会社にしかないと思うが、月額40万ぐらいのレンタル料がかかるが、ニューヨークの図書館に行けば、5台、6台ずらっと並んでいて、ホームレスでもただで使える。だから、そういうビジネスモデルを考えていく必要があると思っている。

例えば日経テレコンなんかは、日本でも、公共図書館、400館ぐらい導入して、これはいわゆる従量制ではない定額制のプライスで、かなり安い値段で使えるという状況があるので、そういうビジネスモデルを考えていく必要があるのではないか。(常世田委員)

- アカデミックな方たちは、基本的に何らかの研究ライブラリーのところはかなり近い人たちがいるのではないかと。というのは、公共図書館は横で見ていると、対コスト効果と利用者みたいな議論がいつも出てしまう。その上で、では出版界も何か提供したいということで、もちろんそれはビジネスの領域としてトライアルしてはいるが、例えばジャパンレッジなど、実際これは完全にいろんな出版社さんがもう既に図書館に入っており、例えば吉川弘文館さんの国史大辞典の巻がなくなっており、図書館が傷んでしまったのでそれだけ買いたいといっても、買えない状況があったものに対して、その全巻を例えばデジタル化する。あるいは、平凡社さんの民俗歴史体系、東洋文庫も、欠本だらけだったものが、今は全部デジタル化している。

大学図書館はかなり入っているが、公共図書館はまだほんとうにこれから。これはまず学術研究論文の先に、この辺は絶対入るといいだろうと思うものはまだこれからなので、それはニワトリ・卵の問題、お金がある、ないの話はあまり持ち込みたくないが、そういうものこそちゃんと埋めて、利用していただかないと、逆に言うと、結局ビジネスもショートしてしまう。今はかなり多くの出版社さんの持ち出しと赤字の中で、もうおしまいかな、これは無理かなと。利用していただかないことには。

なので、電子書籍は、ほんとうはお金になればすぐにでも動くから放っておいてもいいと思っているが、その手前であるが、出版界がつくらなきゃいけないとわかっているもので、やっぱりやっているところもあるが、ここは全然苦しい。そうすると、その先に、まずデジタルアーカイブのほうとしてやっていただくところはどこなのかなというところは、もう少し限定的にさせていただいたほうがいいかなとは思っている。ちょうどボーダーのところは、今のところ全然、出版界も持ち出しでやっているから、ライブラリーとか、皆さんと一緒に考えて、お金が回って続く方向は考えましょうねというのはある。

(植村委員)

- ・ ビジネスの話になったので、あえて話を戻したいのだが、先ほど杉本先生のほうから、現在の情報環境におけるという話があり、広い意味で情報環境だが、今のインターネット時代における情報の公共性といったようなものが、随分情報の公共性という考え方、変わってきていると思う。それこそGoogleがまさに変えていたりする。本はあまねくアクセスできるべきだとか。実際インターネットの時代になると、データ、情報というのはアクセスできて、価値が出るものというふうなものがあって、考え方も変わりつつある。あるいは公共という考え方、例えばオープンガバメントのような、政府の情報もなるべく公開すべきだと。そういうことがあって、そういった中で、いわゆる知のアーカイブとして、どのぐらい公共性に対してある種のメッセージを出せるかというのも1つ重要ではないか。ここで短絡的にすべての情報は無料で公開すべきだとか、そういうことを主張するわけではなくて、ただ、そういった方向を確認した上で、それぞれの持っているビジネスはそこにどう関連し合うのかということ、そういうのが、単にデジタルアーカイブをつくるビジネスを整理するというようなことを超えて考えておくべき点かなと思います。

それを考えた上で、現状はここはビジネス的に成り立たないとか、そういう問題を見ていくのがいいのかなと思う。でないと、どうしてもお金がないよという話になってしまうと、そこで話がとまってしまう。実際にお金がないよと言いつつも、じゃあ、Googleがやるからいいんですかと言われると、話が振り出しに戻ってしまう。じゃあ、Googleが全部本のアーカイブの費用は出しますから、それならいいんですかとか、そういう議論をここでしたいわけじゃないんじゃないかなというのが、今の議論で危惧したところ。(武田座長代理)

- ・ デジタルアーカイブのメンバープレーヤーMLAにおいて、きっと人、お金があればやりたいよということは、多分そんなに異論はない。

その一方で、実際に連携をしようといったときに、いろんな課題がありますねということ。特に、MLAの連携は、ある種のデジタルアーカイブ関連系という話になるかと思うが、連携することによってどういう付加価値が生まれるのかということに関して、何かあるというふうに私は期待したい。(杉本座長)

- ・ この検討アジェンダのつくり方とも関係するが、このアジェンダを見ると、検討の目的というところで、デジタルアーカイブ間の相互連携ということが書かれていて、その

後、ずっと各館種ごと、各館種においてという連携を書かれている。確かにデジタルアーカイブは、単につくるだけではおもしろみがなくて、例えば各館種ごと、LLとか、AAとか、MMとか、同じ館種ごとに連携することによってデジタルアーカイブの可能性は広がると思う。それはいいが、次のステップとして、MLAとか、違う館種を連携することによって、さらに相互連携というか、デジタルアーカイブそのものの可能性が広がると考えると、1ページ目は、AAとか、LLとか、MMの連携について書いているが、2ページ、3ページになって、またデジタルアーカイブの相互連携というふうには書かれているが、それは1ページの各館種ごとの連携を踏まえて、そして次のステップとして、異なる館種のMLA、MLとか、LAとか、LMとか、そういう異なる館種の連携をすることによってさらに一層デジタルアーカイブの価値は広がるんだということをもう少し明示したほうがいいのではないかな。すなわち、デジタルアーカイブの連携というのは、まず個々のアーカイブの構築があり、そして館種ごとの連携があり、次のステップとして、MLA、あるいはもっと違うところと異なる領域との連携という、2段階、あるいは3段階のステップをもう少し明確に書いたほうが、今までのデジタルアーカイブ構築とは違う局面に入ったんだなということがメッセージとして伝わるのではないかと。そこまでいかないと、デジタルアーカイブのおもしろさというのは、なかなかメッセージとして伝わらないのではないかな。

例えばすごく単純なことを言えば、私の美術館に来て、所蔵作品を見て、絵を見る。さらにその作家、さらにその作品について、あるいはさらに周辺状況について知りたい、あるいは知らせることができるようにするためには、MとLがつながっていけば、私の美術館には展示場があって、作品があるが、フロアは違うけれども、ライブラリーがある。MからLへの動線が引ける。あるいは、お隣に公文書館がある。そういう可能性は幾らでもあるということ。(水谷委員)

- 電子になって、アクセスの制限が逆に入ってくることもある。電子にするコストを持った人は、電子を抱えようとするから、単純にオープンにしようとしなくて、紙だったら流通していたものが、電子になったら流通しなくなる可能性がとて多く、そうすると、紙で、普通で言えば図書館間の連携って、美術館、博物館にも図書室があり、いわゆる紙でのILLがあったが、これが電子になったときどうなるかという、逆の意味での情報的な閉鎖感実は出てくる。逆の意味で、それをちゃんとやっていかないと、情報流通ができなくなっちゃうんじゃないかという危機感も一方である。(入江委員)

- ・ 私は今の座長先生のお話でいうと、それはG o o g l eでいいよねという人。つまり、そんなことはわからないと。要するに、データをどう組み合わせれば役に立つなんて、やる前にわかるわけがない。だから、今の段階で必要なのは、コーパスというか、とにかく数を**10倍**、**100倍**に増やすことだけ考えておけば今はいいんだと。組み合わせはみんながやってくれる。だから、あふれ返るような状況をつくるのがまず大事だと。**10倍**、**100倍**にするために何を考えるかということならいいけれども、資源共有化の根本的な弱点はそこにある。結局データを集めてきても、キーワードは統一されない。つまり、分野が違えば全然キーワードが違う。あと、連想検索みたいなことを考えるべきだが、連想検索を考えようとする、今では母数が足りない。もう少し母数が上がってこない、例えば機械的に連結関係だとか、近い、遠い関係をはかって、いいのが出てくるなんていうことをやろうとすれば、コーパスにならなきゃいけない。そうすると、今の規模では小さ過ぎる。だから、**10倍**、**100倍**にするための人のかけ方、お金のかけ方であり、もっと言うと、お金をかければできるということは全然なくて、お金をもらっても、ものばかりしか買えないので、全然役に立たない。機械を買って終わり。機械は5年たったらつぶれる。そうでなく、**10年**雇える人。それも、ドクターというよりは、マスターぐらいの人を**10年**かけてその筋の技術者に育てるようなお金のつけ方。そういうことがあるのなら、お金をつけて、人をつけてもらってという議論は成り立つのではないか。予算システムに根本の原因があるが、要は、人が雇えない。雇う人は全部一時雇用の人。国立歴史民俗博物館でも働いている人はみんなアルバイトで、自給940円で交通費つかない中で何年も何年も来ている。最近では雇い過ぎるといけないから、ほかのところに移りなさいなんていうご指導も丁寧に入ったりする。

そんな中で、知というのは積み重なっているイメージ、どこかに積んであるもの、うず高く積まれていくもの。デジタル以前はこれを探すということが非常に大変で、専門家が探す検索の役を果たしていた。だから、専門家が必要だった時代があった。ところが、G o o g l eみたいなのがやってきて、それはコンピューターで探せるんじゃないかと。そうすると、何が起こるかという、人間、探せるようになりますと、それを積んでいる人間がいることを忘れて、いずれ何か役に立たないがらくたばかりあるんだけど、だれがこれ責任を持って面倒見ているといたら、だれもいないという状況がこのままだと必ず来る。つまり、こういうところでもし議論していただくなら、それを積む役目の人。探すことが便利になるのは構わないが、積む役目の人をどうやって育てるの、

どうやってお金積むのという話をぜひやっていただきたい。それは、先ほどの出版の話にしても、書き手の供給の問題につながっているはず。だれが書き手になるのかと。Googleが書き手になってくれるというのなら、それでもいいんですよ。だから、Googleがやっちゃったことというのはそういうことですよね。書き手になれる人となれない人というか、なれない人はGoogleでいいじゃんという話になるから、だんだんそういう淘汰を我々に突きつけているんだと思う。(鈴木委員(安達構成員代理))

(6) ○「デジタルアーカイブの構築・統合に関するガイドライン素案」について松田統括補佐より、説明があった。

- ・ 前回の会議でもご協力をお願いしているとおおり、総務省において、昨年度からデジタルアーカイブに関する技術的な課題について調査を行っている。今年度の内容としては、昨年度、仮に取りまとめているデジタルアーカイブ構築のガイドラインについて粗々の素案であるが、これについて、しっかりと修正をかけていきたいと考えているところ。
- ・ また、構築だけではなくて、相互連携、統合的な利用のためのアンケート調査も行い、ガイドラインの素案を何とかまとめていきたいと考えている。

お手元の資料知2-5と2-6というものが、それぞれ構築と統合といったものになっており、これについては、あくまで素案であり、今年の6月ぐらいを目途として、皆様のご意見、あるいは世の中の人の意見も踏まえて、ガイドライン案としてまとめていきたいと考えている。

構成員の皆様には、メール会議も活用しながら、ガイドライン案についてご意見等ちょうだいしたい。

○また、資料知2-5「デジタルアーカイブの構築のためのガイドライン素案」、資料知2-6「デジタルアーカイブの統合のためのガイドライン素案(骨子)」について、インフォコム株式会社より説明があった。

- ・ 資料知2-5からご説明させていただく。資料知2-5については、「デジタルアーカイブの構築のためのガイドライン素案」ということで、これは昨年度につくった報告書を前回の会議の中で概要を説明させていただいた。これについて、今期はデジタルアーカイブの保存、整理、提供といった形で、より業務フロー上の中でデジタル

媒体を円滑に扱うといったガイドラインを提供するためのAPI、システムのユーザーインターフェース、バックアップ保存等について、幾つかのポイントを挙げさせていただき、資料知2-5の別紙にその素案のコメントを書かせていただいている。これはたたき台と考えているので、会議の場やメーリング会議の中でいろいろ先生方のほうにご意見をいただきたい。

- ・ 続いて資料知2-6について。最終的に「デジタルアーカイブの統合のためのガイドライン素案」というのを確立していきたい。まず、背景、方針、それから統合のためのガイドライン策定のための調査(骨子)、それから、統合のためのガイドライン方向性についての総括、それから議事録ということで、これが最終的な報告書の形になるようなイメージでとらえている。

まず、背景、方針というところで、背景はデジタルアーカイブ間の相互連携が可能となる技術、標準化の推進。方針については、地域毎、規模毎の現状の資料毎の保存、所蔵の現状に即したデジタルアーカイブの統合ということを方針に置いている。

その中で、私どものほうは、2500機関のMLAへのアンケートをとって約40%の回収率で最終報告書のほうはまとめさせていただいている。

その結果の概要について、現状、約100%のメタデータ整備が行われているというのが、約3割の機関が実際に取り扱っているという回答。逆を言うならば、70%はまだまだやりたいデータが残っている、もしくは全くデジタル化されてないとか、そういったところに結びつく。

- ・ 独自資料についてのメタデータ準拠している目録規則や、メタデータ構造に関しての確認をさせていただいたが、その中では、独自資料が多いということで、エクセル等を使った管理にとどまっているというところが結構多かった。

5ページ目のアンケート結果の考察、メタデータを共有という観点からの報告について。こちらのほうは、実際に全国的な連携よりも、先ほど先生方からも御意見が出たところであるが、まずは地域、同館など、近場からの連携が望ましいんじゃないかという点も、各MLAからも意見がでてきている。

また、次はヒアリングというのを実施中。沖縄、都内、大阪を今後回る予定で、MLAのアンケートに基づくポイントを確認していき、ご報告したい。

さらに8ページ目のとおり、38社のベンダー向けのアンケートを考えているところ。クラウド化の対応も加味した形で、アンケート結果を出させていただく。9ページ目

の実証実験について。2月下旬から3月の中旬にかけて2、3機関に実際に行き、エクセルベースのデータになっているものについて、いかにそこからデータ登録をして、公開まで持っていけるような要素があるのかということについて予定している。

これら調査研究、実証実験を踏まえて、報告書まとめてまいりたい。

(これについて、以下のやりとりがあった。)

- ・ 資料にあるように、最も普及しているのは、一般的なのはUTF-8なのか。(安達委員代理(鈴木))
- ・ 図書館のシステムが使われている形は、UTF-8がベースになっており、一番当たるかなと考えている。ただ、デジタルアーカイブを管理する上で、今回の調査でもわかったことだが、現状は、SJIS、エクセルで管理されているということが圧倒多数に多かったので、そういったところは今後の調査の上で、報告したい。(インフォコム(栗原))
- ・ ワード、ウィンドウズで、ユニコード形式のテキストをつくるならば、UTF-16のリトルエンディアンでBOM使っており、メールでエンコーディングする以外、UTF-8って使っていないのではないか。(安達委員代理(鈴木))
- ・ 保存のことについて記述がかなり不十分ここは改めて大きくするなりして、ぜひ別項で書いていただきたい。(山崎委員)
- ・ 本日も提示した構築のためのガイドライン素案、現状進行中の統合のためのガイドライン素案(骨子)の調査の状況については、今回、資料を提供させていただいたが、メール会議で引き続きご意見をいただきたい。昨年度から、このガイドライン素案を調査の中でつくってきたが、国会図書館、杉本座長からお話をお伺いして、国会図書館のものがベースになっているという状況にあるため、これをより汎用的、もっと役に立つようにするのであれば、構成員のご知見を入れていただきたい。(松田情報流通振興課統括補佐)

(以上)