

第1章

基盤機関アーカイブズ共有化の現状と今後の課題

五島 敏芳

国文学研究資料館 助手

安倍 尚紀

総研大葉山高等研究センター 上級研究員

1. 基盤機関アーカイブズ共有化の現状(五島)

基盤機関アーカイブズ共有化の現状について報告させていただきます。今回の報告は要約すれば、以下の3点にまとめられます。

1. 既知の各基盤機関収蔵資料の検索手段(資料目録)の情報について、共有するための技術的環境が整った。
2. ただし、その条件に合わせた記述ないし目録の作成・変換、周辺管理的情報の自覚的整備が必要である。
3. そのため、ときに資料収集受入・調査の作業段階に遡って、それらの手法の変更が求められる。

以下、それぞれについて簡単に要点を述べておきます。

1.1. 整備された技術的環境

技術的環境については、特定のアプリケーションソフトに制約されないXMLテキストファイルの国際的規格を採用しました。アーカイブズの電子的検索手段(オンライン目録)のデファクト国際規格にはEAD(Encoded Archival Description=符号化記録史料記述)があり、日本でも、それをもとにした検索システムの利用が十分可能であることがわかってきまし

た。このシステムの最大のメリットは、国際規格であるために、海外のアーカイブズ機関との情報交換・共有が可能である点です。

私の所属している国文学研究資料館（国文研、NIJL）で、この研究や実験を進めてきましたが、このデータ形式は、EAD 発祥の地、カリフォルニア大学バークレー校のものとはほぼ同じデータ形式です。国立公文書館でも最近 EAD 形式を使っていて、イギリスの国立文書館のデータ形式に類似はしていますが、データ形式に一部ルール違反がありますので、そのまま新たに採用されることは避けたほうが賢明です。

一方、EAD 形式の不利な点は、データ構築の手法が未確立であることです。ツールとしては、国文研「史料情報共有化データベース」がありますが、現在まだ種々の問題を抱えています。

1.2. 目録データの作成

目録データの作成にあたっては、まず資料群（記録群、コレクション）としての把握を念頭においてください。アーカイブズには、その元となる記録書類を作成し授受し保存した（＝管理した）出所（団体・個人・家にかかわらず）があるか、または収集者がいるわけですから、その資料群全体のだいたいの内容、出所の履歴、使用上の注意などを意識しておかなければ EAD 形式のデータ作成はできません。ですから、まず資料群全体で、ダンボール箱がいくつあるとか、いつ受け入れたかなどのデータをぜひ作成していただきたいのです。すでにある 1 点ないし 1 件ごとの資料の書誌データは、EAD 要素へマッピングすることで対応できます。

1.3. 周辺管理的情報の整備

まず、それぞれの基盤機関が所有する収蔵資料の一覧を作成する必要があります。利用者は、この一覧から誘導されて特定の資料の情報に行き着くことになるので、EAD ではそういう使われ方を前提とした作り方をしています。総研大でのアーカイブズセンター構想が実現すれば、アーカイブ

ズのバーチャルコレクションができると思いますが、そこでは、まず各機関の資料一覧リストを仮想的に保有する必要があります。同時に、各機関がそれぞれの資料一覧リストをもつことが必要です。それはユーザにとって最初のとっかかりになると思います。それぞれの基盤機関の目録における資料の並び方は、資料群の五十音順、アルファベット順、地域順など、各機関の特性に応じたものでかまわないでしょう。ただし、その場合、機関や資料群の読み仮名、英文名またはローマ字、関係地名など、EADにはないデータ要素も必要になることがあります。

最後に、既存の管理過程もいくらか改善する必要があります。1点ごとの資料整理をしているのであれば、資料群として再整理するためのデータ把握をしていかなければなりません。歴史資料の調査手法に関しては、資料群の書誌データ・管理データのためのマニュアル類¹があり、資料を群としてどう把握するかについて記述してあります。実際、資料がどの棚にどのように並べてあったかなどもアーカイブズとして意味がある場合があるので、それを写真に撮っておくなどの現状記録方法も参考にしていたいただければと思います。

2. デジタルアーカイブズ化の現状(安倍)

まず初めに、昨年度の達成状況と今年度の課題・目標について整理しておきます。

2006年時点(4月24日、25日、全体会議時点)の到達点は下記のとおりです。

¹ 自治体史編纂や「～史料調査会」といった団体の調査報告書に収録されていて、その内容も実例として参考になる。たとえば、つぎの2つをあげる。1)牛久市史編さん委員会・近世史部会。『牛久市小坂・斎藤家文書概要調査報告書』。茨城、牛久市、1993年、122p。2)甲州史料調査会。『山梨県山梨市下神内川区有文書調査報告書』。東京、甲州史料調査会、2003年、178p。

- ・ ウェブページの情報発信
- ・ wiki や xoops などの CMS
- ・ CSV ファイルによるテキスト検索
- ・ クロスサイト検索（該当ページをロボットで巡回収集、インデックス化）
- ・ EAD の基礎理解
- ・ Flash による公開用インターフェース

一言で言えば、昨年の時点では技術の確立が不十分でしたが、今年は、技術的な内容はクリアできたと言えます。このことを前提にして、総研大プロジェクトの全般的な概要とアーカイブズ共有化の現状について報告させていただきます。

2.1. 「共同利用機関の歴史」プロジェクトの特質

総合研究大学院大学・葉山高等研究センターにおいて進行している本プロジェクトの正式名称は、「大学共同利用機関の成立に関する歴史資料の蒐集とわが国における巨大科学の成立史に関する研究」です。

「共同利用機関の歴史」と通称されるこの長い呼び名が前段と後段に分かれていることからわかるとおり、プロジェクトの特色の1つは、①大学共同利用機関の歴史を研究するなかで、同時に、②そのための歴史的史料を収集・管理するシステム（＝各大学共同利用機関および総研大のアーカイブズ）の構築を促しつつ、これらの情報を共有するという実践的側面にあります。歴史研究とそのための資料システム構築の実践がリンクしていることから、両者は車の両輪と言えますが、ここでは、プロジェクト内で「情報共有化」として括られ、共有化セクションとして活動している②に焦点を当ててお話ししたいと思います。

2.2. 情報共有化の射程

まず、「何を対象にしているのか」「どのようなニーズに応じているのか」

「どんな技術を使うのか」といった、情報共有化のコンセプトを明確にしておく必要があります。

●何を対象にしているのか

総研大は、自らをも含む各研究機関にアーカイブズ（史料館）が必要であると考えています。最終的には、研究を進めるなかで、その設置・実践活動を支援し、資料情報の共有化をはかるために、総研大内のアーカイブズ史料の統一的な検索可能「システム」を構築することが目標です。

●どのようなニーズに応じているのか

これについては「2.3. アーカイブズ一般と扱う資料の特質」で述べます。

●どんな技術を使うのか

セキュリティ面での保障、各アーカイブズ現場での制御・更新の容易さなどが考えられますが、それについては、「2.4. 目録の EAD 化（電子的検索手段へ）」以降で説明していきますので、ここでは、以下の点を指摘するにとどめたいと思います。

- ・ 共同利用機関での試みは、研究機関によるいわば「inhouse の」アーカイブズであり、日本国内では先進的な実践である
- ・ アーカイブズの資料が広く活用されるために、「検索手段 Finding Aids」と呼ばれる資料の目録情報（メタデータ）をデータベース化して用意し、この手法として EAD を採用する
- ・ 総研大内部のみならず、外部との連携（国際標準）をめざす

2.3. アーカイブズ一般と扱う資料の特質

アーカイブズとは、「個人・組織がその活動のなかで作成・收受・蓄積した資料で、さまざまな利用価値により保存・管理されるもの」とであると定義されています（五島敏芳、2003、「『国際標準：記録史料記述の一般原

則；ISAD (G)』とXMLの利用」²⁾。

「蓄積」とあることから予想できるように、アーカイブズは、およそ〈群〉として存在する。アーカイブズを構成する一つひとつの記録・文書等の資料は、発生母体たる個人・組織の活動やそれに基づく何らかの体系的な秩序（脈絡）のなかで位置づけられ、脈絡の説明があってはじめて利用できる（五島 2003）。

上記の点が、EAD の採用につながってくる大きな理由でもあります。たとえば、アーカイブズの中に、研究者が誰かにあてた手紙があったとします。発信地を示す消印、封筒の種類、誰から誰にあてたものか、いつ・どんな状況で書かれたのか等々の周辺の情報を把握することによって、研究者自身や、研究の遂行などを知る上で重要な手がかりになります。

この点が、アーカイブズと同一視されがちな「図書」目録との大きな違いです。つまり、アーカイブズ資料は、図書と違って世界に1つしかないオリジナルな活動の痕跡であり、なかには進行中の記録として現在も継続しているものも存在します。したがって、コンピュータによる検索も、図書と違って、1つのファイルを見るだけでは不十分なので、そこに脈絡が必要になってきます。

このために、記述の標準化が不可欠となります。ISAD (G) (General International Standard Archival description=記録史料記述の一般原則)はそのために考案された標準です。ISAD (G) は、正確には規格ではなく、あくまで全般的な標準であり、5年ごとの見直しが提案されています。事実 1994 年に採択された初版が見直されて、2000 年に第2版が公表され、下記を段階的に実現しようとしています。

²⁾ 科学研究費補助金研究成果報告書「国際コラボレーションによる日本文学研究資料情報の組織化と発信」（研究代表者：安永尚志、2003年）に収録、もと2001年10月27日発表の口頭報告。

- a) 一貫して適切かつわかりやすい記述を保証する
- b) 記録史料に関する情報の検索・交換の便宜を図る
- c) オーソリティ・データの共有化を可能にする

アーカイブズには、記録文書をはじめ、オーラルヒストリー資料（トランスクリプト等）をはじめ、各種デジタル資料、映像資料・動画など、さまざまな資料が含まれます。この点も、アーカイブズの大きな特質です。これらをかんがみ、本プロジェクトでは、ISAD (G) に日本のアーカイブズの記述を適用し、その国際標準に準拠した記述を具体的に蓄積していく手段として、EAD/XML を利用することに決定しています。

2.4. 目録の EAD 化（電子的検索手段へ）

EAD に関連する内容については、予算をかけて業者にシステム開発を外部委託する、という方法もないわけではありません。しかしプロジェクトは、幸いにも五島先生を中心とする国文学研究資料館の協力を得ることができ、アーカイブズ基礎論を含め、技術的な情報面で多大なサポートをしていただきました。したがって、EAD は(すでに国文研で運用されているように)技術的には完成しています。

残された問題は、先行研究もかつて苦労したような実際の記述についてだと言えます。これまでのところ、人員や技術面での理由から資料の蓄積が滞っていますが、現在は通常のデータベースを変換するツールなど、さまざまな自動化の手段を模索していますので、近いうちに解決できると期待しています。

2.5. 達成点と今後の課題

冒頭に指摘したように、2006 年度は、アーカイブズ共有化の技術的環境の整備に着手し、かなり達成しました。

また 5 月 31 日、6 月 1 日に、国文学研究資料館 (NIJL) と直接打ち合わ

せた結果、総研大プロジェクト独自の開発路線から、NIJL データベースに同期させてもらう方法に路線変更することとし、7月には史料情報共有化データベースに登録されました。これによって、技術的な問題はかなり解決しました。

さらに、かつてはこれら「情報共有化」セクション全体の会議は、研究会などの機会に便乗して開催していましたが、2006年8月から、通称「研究機関資料目録EAD化会議」という、情報共有化の会議を毎月1回のペースで定期的に行うようになりました。この会議では、分子科学研究所の木村克美名誉教授、葉山へ着任した伊藤憲二助教授の協力を得ています。

【表1】クロスワーク表の作成も、この会議の成果の1つです。一番左側の列はEAD項目で、核融合研と高エネ研が独自に作成している項目がどこに対応するのか、どの部分が欠けているのかが一目で分かります。これによって、各基盤機関のデータベースの同期化が可能になりました。

【表1】クロスワーク表

2006-08-04ワークショップ時作成。

登録先(必須)	[NIFS-archives]	[KEK-archives]
公開・非公開	-	-
識別記号(必須)	[[2103-746271]+(コレクション	[[0803-481173]+コレクション
史料群記号(必須)	(IDNum)	Identifier
名称(必須)	文書資料名、副題	title, -en
年代 上限(西暦)	時期(始期)	
年代 下限(西暦)	時期(終期)	
主年代		
年代注記	発行・作成年月日	Date
記述レビュー		
規模(書架延長:数量)	-	units
数量媒体等詳細	資料の形態	Format
出所・作成	資料作成者・機関	Creator
履歴		
関係地		
領主(→主題subject)	キーワード	subject, -en
入手源	資料提供者	Contributor
構造と内容(範囲と内容)	資料内容	Description
追加受入情報		
公開条件	-	rights
複写条件		
使用言語	-	Language
関連史料(→関連資料)		
出版物	-	Relation

その他、核融合研、高エネ研、分子研、その他科研費プロジェクトなどとも連携・協働して、さまざまな実験、研究を進めています。また後ほどデモンストレーションをごらんいただきますが、インフラ上は、インタビュー等、動画の公開も可能になっています。

3. 史料情報共有化データベースのデモンストレーション(五島)

史料情報共有化データベースは、資料目録全部を共有するわけではないのですが、どういう資料群なのかについての情報を全国の機関で共有するためのデータベースです。この資料の概要を入力する編集機能について、デモンストレーションでご紹介します。

Web 入力画面に、資料の全体のまとまりを階層的に登録することができます。この階層的に登録できるという点が、図書資料との大きな違いです。図書資料は、雑誌、書籍など1つ1つをすべてフラットにデータ化していきますので、検索しても階層性はまったく分かりません。もともと階層性を意識してつくられてはいないわけです。それに対してアーカイブズでは、組織、個人の記録を階層的につくっていきます。それをしないと、何の資料か分からなくなってしまいます。

たとえば、ある研究会を開催したときに購入したお茶などの領収書が、研究会の議事録などと一緒に綴じ込まれていたとしたら、それが検索結果で単独で出てきても、何のことか分かりません。それを理解するためには、研究会で用意されたということが分かる階層的記述が必要なのです。

そして、階層的検索結果をそのままビジュアルに見せる検索システムが可能になりました。もちろんキーワードによる横断的検索も可能です。

国文研の予算もかなり厳しい状況ですので、できれば予備の検索システムを稼働させるか、あるいは基盤機関だけを対象に稼働する検索システムをもったほうがいいが安全だと思います。総研大が検索サービスの中核になっていただけることを期待しています。

〈質疑応答〉

—— デモンストレーションでの検索が一番深いところまでいくと、どうなるのですか？ 最終的には1件の情報になり、その情報が必要な場合は、保管している研究所などに問い合わせる、ということですか。

五島 そのとおりですが、最終的にいきついた1件に画像や本文のようなデジタル情報があれば画面上に表示可能です。リンクをはらないようにして、デジタル情報にたどりつかないようにすることも可能です。ちなみに当館の場合、リンクから他のデジタル情報へジャンプする仕組みは、五島または国文研から無償提供できます。資料目録の情報をナビゲートする仕組みも無償提供できますが、いまのところ（横断的検索を含む）1件ごとのキーワード検索の仕組みには商用ソフトウェアを使用していて、この購入にお金がかかることが一番の問題です。

—— 高エネルギー研、核融合研、国文研で、それぞれのアーカイブズのデータを検索する共通の方法として EAD 方式に決めたということですね。別の方式の可能性はないのでしょうか？ というのは、高エネルギー研は以前から独自の方法でアーカイブズの整理を始めていたと思うのですが、今回の EAD 方式に変更して、使いやすいと感じられているのか、どうか。他の機関の今後のこともあり、その点が気になります。

安倍 1つ考えておかなければならないのは、EAD は世界標準である、ということです。ですから、EAD の形式を整えておけば、他の研究所とも同期がとれるのが最大のメリットだと思います。

—— 高エネルギー研、核融合研と相談し、協力を依頼した上で、EAD による作業を進めています。カリフォルニア大の膨大なアーカイブズも EAD でつくられており、われわれのプロジェクトも登録される話も進んでいます。そういう意味で、もはや世界的にデビューしたので、今から引き返し不能というところもあります。

- 2006年4月の総研大でのワークショップでEAD移行に決め、核融合研でも、EADをベースにして作業を進めていると理解しています。
- 五島 アーカイブズの世界でも国際標準を採用し、情報を共有、交換しようという世界的流れになっています。その流れにのっとってデータを作成、管理するほうが、アーカイブズの実物管理にも非常に効率がよいのです。何千点にものぼる資料を、1件1件検索しなければ情報が共有できない、といったことのないように概要を把握するための枠組みでもありますので、ぜひ使っていただきたいと思います。また絶対満足いただけるものと確信しています。
- 安倍 これから実際に記述作業を進めなければならないという点だけが、問題点として残されています。
- これは、世界標準だとすると、すべて英語なのですか。
- 日本語システムでEADを使うことは大変でした。自分たちですべて作り直すのなら大変だと思ったのですが、五島先生がすでに着手されていたのでぜひ活用させていただくことにしました。

(補記) 国文学研究資料館「史料情報共有化データベース」は、つぎのURLを参照のこと。<<http://basel.nijl.ac.jp/~isad/>> 同データベースは、どなたでも自由に閲覧できますが、編集機能の利用には編集ユーザとしての登録が必要です。同館へお問い合わせください。また、EAD-XML検索システム(国文学研究資料館サイト)は、つぎのURLを参照のこと。<<http://basel.nijl.ac.jp/~eadfa/>> 資料群の概要から1点にいたるまでの検索が可能です。