

湯川秀樹・朝永振一郎・坂田昌一各記念史料の整理とデータベース化 および オープンソースアーカイブズ情報システムによる公開

高岩義信†, 五島敏芳‡

† 高エネルギー加速器研究機構 (情報資料室 史料室)

‡ 京都大学総合博物館 (京都大学研究資源アーカイブ担当)

物理学研究者である湯川秀樹 (1907~1981), 朝永振一郎 (1906~1979), および坂田昌一 (1911~1970) の記録書類・研究資料等を整理しそれ利用する科学史等の研究を推進することを目的とする研究が科学研究費助成金 (科研費) によって実施された。2008 年度から 2010 年度まで (第 I 期) の課題「湯川・朝永・坂田記念史料の整理と史料記述データベースの整備」とそれに続く 2011 年度から 2013 年度まで (第 II 期) の課題「湯川・朝永・坂田記念史料から分析する日本の素粒子研究者の系譜」である [1]。本報告はその活動のうちアーカイブズ学にかかわる話題について報告するものである。

1. 湯川秀樹・朝永振一郎・坂田昌一とその記念史料

湯川, 朝永, 坂田は日本の原子核・素粒子理論の礎を築いた物理学者であり, 多くの研究者がそのもとで育った。この三人の記録書類・研究資料等が記念史料として, 京都大学基礎物理学研究所 湯川記念館史料室 (湯川記念館史料室, 1979 年設立), 筑波大学 朝永振一郎名誉教授記念室 (朝永記念室, 1983 年設立), 名古屋大学物理学教室 坂田昌一記念史料室 (坂田記念史料室, 1973 年設立) に保存されている。これらの記念史料室が設置されるようになった経緯はそれぞれ異なる。湯川の場合は日米の研究者による日本の素粒子物理学の歴史に関する共同研究の際に科学史研究者を含む米国の参加者から保存の重要性の指摘を受けた物理学者たちによって, 海外の物理学関係の類似の研究者の資料保存の施設・アーカイブズから手法を学びつつ, 資料の管理・維持を続けてきた。朝永の場合はその没後, 遺族から大学に資料が寄贈されたことをきっかけとする。また, 坂田は在職中に亡くなったが, その指導の下にあった研究者が著作集編纂のために収集した資料を保存する場を大学に設けたことに始まる。いずれも湯川の記念史料の整理の進め方に学びつつ整理作業はそれぞれほぼ独立に行われていた。どの記念史料にあっても専任のアーキビストがいるわけではなかった [2]。時の経過とともにそれぞれの施設で精力的に活動していた者の転出・高齢化・物故により, 活動が停滞して各史料室の状況を把握しづらくなってきていたが, これらの物理学者の生誕百年記念事業が企画されるにあたって, その重要性を共有し各記念史料に注目する者たちが各史料室の関係者と協力し, 資料管理を維持していくことが模索された。そのためのプロジェクトが科研費の助成を受けた共同研究として実現した。

2. 史料整理における課題

各史料室の抱えていた課題の第一は, 部分的な整理と資料目録作成はされていたものの, 未整理の資料が多いことだった。資料整理をすすめ検索手段 (資料目録) を整備することが急務であった。また既存の資料目録の形式は各史料室で異なっていた。交流も密であったこの三名をとりまく歴史的状況の理解を深めるには, 各史料室の資料検索を共通の環境で横断的に行えることが望ましいという観点から, 資料の整理・目録整備を協働して進め, 目録のおよその形式の共有をはかることになった。

2.1. 資料の編成と階層性の把握

そのためアーカイブズの世界の原則や標準類に則って作業を進めることにした。具体的には, 資料の編成を特徴づける階層性を尊重して整理し, 資料の概要を大枠でとらえてから徐々に詳細の整理と記述にかかるようにしたこと, つまり大量の資料が存在する場合, 資料 1 点ごとの細部

の整理よりもキャビネットや箱などの上位階層の概要の把握を優先し、詳細の整理や記述データ作成には優先順位を付けて進めた。その際、既存の資料目録は尊重し（さかのぼっての再整理はせず）その記述データを利用した。

2.2. データベース化とメタデータ標準の採用

検索手段を提供するためのデータベース化にあたっては、このような段階的・階層的に資料群を把握するのに、階層的に記述データを符号化できる電子的検索手段のデファクト標準 EAD, *Encoded Archival Description* が適している。はじめ、EAD を実装していた国文学研究資料館「史料情報共有化データベース」（2008 年当時）を採用し、記述データの作成と 3 史料室間での共有を同時に進めようとしたが、一括のデータ処理や細かな使い勝手の改善に限界があったため、市販データベース管理ソフトウェア（FileMakerPro）によるデータベース構築が試みられ、EAD データの出力機能原型も準備して各史料室に導入した。あるいは階層性の構造が複雑でない場合は簡便な EXCEL で基本的なメタデータを蓄積していく方法がとられ、必要に応じて EAD データに変換できる FileMakerPro のフォーマットにインポートして処理する方法も取られた。それぞれの記念史料のデータベースは共通の標準に準拠した形式に変換できるようにして、独立に管理される。

3. 史料の活用に向けた環境整備

3.1. 検索手段の公開

第 II 期では記念史料にある記録資料データの活用に向けての環境整備と実際のケーススタディを第 I 期の基礎の上にたって行ってみることを目標とした。保存史料を活用するには資料情報の公開が必要である。第 I 期では検索手段を公開するにあたって、やはり標準類のデータを取り扱える汎用のオープンソフトウェアによる Archon[3]を採用して Web サーバを構築した。各史料室で作成したデータベースからエクスポートされるデータをサーバに投入する手順の確立を行った。これに関しては今後の展望もふくめ本報告の後半で詳しくふれる。

Web サーバからデジタル化されたコンテンツデータにアクセスできるようにすることに関して、技術的な課題を別として、公開が適切な資料とそうでないものの区別、公開を許可する利用者の認定などの課題の検討が必要である。そのための方針の基準とメカニズムの設定にまだ検討の余地があると思われる。

3.2. 典拠レコード（コンテキスト・データ）のデータベース化の課題

資料を利用した科学史研究を第 II 期の目標に加え、日本の素粒子物理学者の系譜の研究をテーマとして設定した。その目的に沿って史料に現れる人名や組織に関するデータを蓄積する「典拠レコード」データベース整備を文書資料のデータベースと連携をとる形で整備するという課題を検討したが、資料管理のためのデータとして想定された標準のデータの構造を史料に基づく歴史分析に適したものとして利用することの間にまだ距離がある。このデータベースを構築可能に見える既存の Archon のようなシステムでどのようにこれを利用できるか、実際のところまだ検討が十分ではない。

4. オープンソースアーカイブズ情報システム Archon

第 I 期・第 II 期とも経費的制度的に限界があったため、アーカイブズ情報システムのうち商用のそれを導入することは難しかった。史料の情報の活用、公開には、使用の自由度の高いオープンソースシステムを選ぶ方向となった。

4.1. Archon の選択

第 I 期当時、オープンソースアーカイブズ情報システムとしては、4 つの選択肢があった。CollectiveAccess (<http://www.collectiveaccess.org/>)、Archivists' Toolkit (AT, <http://www.>

archiviststoolkit.org/) , Archon, ICA AtoM (https://www.ica-atom.org/) である。

アーカイブズ情報システムを比較検討した報告書[3]の計 64 項目の状況を点数化すると、ICA-AtoM は、もっとも点数が高かったものの報告書当時に実装のない項目が多く、不十分だった。AT は、全般にもっとも優れていたが、Web 公開機能がない。Archon は、業務システムも Web 公開機能もバランスよく備えていた。

AT は日本でも紹介があり、日本で紹介のない Archon に注目し日本語化に取り組んでいたことから、Archon を採用した。

4.2. Archon の概要

Archon は、もと The Sousa Archives (米国の作曲家・指揮者 John Philip Sousa [1854-1932, マーチ王]のアーカイブズ) で、Web 公開を目指して開発されたという。

Archon は、ブラウザを介して利用するサーバ型システムである。その動作環境は、Linux, Apache, MySQL, PHP (LAMP) で、オープンソースの基本ソフトウェア類である。その内容は、アーカイブ資料管理の業務に対応し、パッケージで構成される。それぞれ、Archon 管理 (システム全体の管理)、コレクション管理、作成者データ管理、主題データ管理の 4 つが必須パッケージで、受入、電子図書館 (デジタルアーカイブ)、視聴覚資料自己評価の 3 つがオプションパッケージである。デフォルト導入設定で、視聴覚資料自己評価をのぞき、すべてインストールされる。

4.3. Archon におけるデータ

各種データを生成・編集は、受入から整理・編成、検索手段構築、電子図書館 (デジタルアーカイブ) による Web 公開まで、アーカイブ資料管理の業務の流れに沿って進めていくことができる。コレクション中の図書資料については書誌データのかたちでも取り扱える (ただしコレクションレベルだけとリンクできる)。

受入に際しては、その時点で必要なコレクション概要のデータを生成できる。つづく整理・編成や記述、検索手段作成に際しては、受入時のコレクション概要のデータをもとに、シリーズ、ファイル、アイテム等の下位レベルのデータを作成できる。Archon では、大きくコレクションのデータと、下位レベルのデータ (CollectionContent) に分けて取り扱う。これらの記述データは、いずれも標準類に対応しているものの、CSV 形式テキストファイルによるインポートに対応している要素セットは限定され、記述レベル深度も基本的に 3 階層に限定されている【後掲表参照】。

Web フォームを使った記述データの入力では、出所等なんらかの (作成者) の情報を入れようとしたとき、その入力欄は自由記述ではない。あらかじめ人名・団体名・家名 (人名等) を入力しておく必要があり、その入力済みの人名等から選択するかたちをとる。

画像・映像・音等のデジタルデータ (デジタルオブジェクト) も登録することもできる。デジタルオブジェクトには、そのための説明記述を入力することができ、このデータと記述データを結び付けることができる。

4.4. Web 公開 (デジタルアーカイブ)

Archon も、たいいていのオンライン資料目録ないしデジタルアーカイブの基本機能を持っている: 1) キーワード検索または各種ブラウザ索引 (アルファベット順), 2) 検索結果一覧または索引抽出一覧, 3) 検索結果詳細表示, (あれば) 4) デジタルオブジェクト表示, が実現する。検索結果一覧等は、アーカイブズの情報検索システムとしての基本提示にしたがい階層的に表現されるが、横向きのツリー状ではなく、画面上の空間を節約したパス状の (所属階層 1→所属階層 2→該当レコードのような) 表示である。

デジタルオブジェクトは、単純表示かダウンロード形式にのみ対応し事実上容量に制限がある。別サーバでの提供等 URL 参照にも対応している。

4.5. Web 公開のための作業

3 つの史料室において FileMakerPro データベースのデータを Archon へインポートする際、インポート形式で使用できる項目と記述レベル深度に限りがあったため、元データベースの項目を統合し割り当て、記述レベルを引き上げるか特定の深度に固定して取り込んだ。また既存資料目録のデータを遡及変換した分については、コレクションの単位ではなく刊行単位で便宜的に処理

している。この成果は、京都大学基礎物理学研究所のサーバより公開されている[5]。
デジタルオブジェクトは、提供のあり方の議論を深めることができず、公開できていない。

5. まとめにかえて

Archon の各機能を試用した結果、3つの史料室の史料の記述情報を連合的に Web 公開することはできた。各史料室がボランティアまたはプロジェクトベースで維持される小規模アーカイブズであること、未整理解消の課題解決の優先といったことから、サーバ型システム導入・維持のハードルが高く、今回は物理学界の全国共同利用研究所のサービスに依存して実現した。

そうした困難はあるものの、Archon のような受入から公開までを支援する業務システムの導入は、日常に更新される資料管理の情報を公開情報へ反映しやすい機動性の利益がある。前述の典拠レコードの蓄積は、Web フォームでの入力への制約に適応し導かれた側面もあれば、今後の各史料室での資料管理における基礎データ整備の意味を持つ。

なお、Archon は AT と統合され、新しく開発された ArchivesSpace として維持されていく。今後 Archon・AT に問題が発生しても対応はなく、システム・データの移行が必要となる。アーカイブズ情報システム導入を推進するとすれば、関係者の積極的関与がもたらされる。

参考文献

- [1] 『湯川秀樹・朝永振一郎・坂田昌一の遺した史料---記念史料の紹介---』(筑波技術大学, 2011) ;
『湯川秀樹・朝永振一郎・坂田昌一の系譜の探究へ向けて---記念史料の整理から活用までの課題---』(筑波技術大学, 2014)
- [2] Prog. Theor. Phys. Supplement, No.105, 1991.
- [3] Archon 公式サイト. URL. <http://www.archon.org/>
- [4] Spiro, Lisa. Archival Management Software: A Report for CLIR. Council on Library and Information Resources, 2009, 119p.
Online, http://clir.org/pubs/reports/spiro/spiro_Jan13.pdf
(Wiki, <http://archivalsoftware.pbworks.com/>)
- [5] <http://yhal.yukawa.kyoto-u.ac.jp/archon/>

表. CollectionContent インポート用 CSV の構成

列	項目名	サンプルデータ	項目名日本語				
1	Collection Identifier	99999	コレクション識別記号	15	Arrangement	ARR FIELD	編成・整理
2	Series #	1	シリーズ番号	16	Biographical/ Historical Note	BIO FIELD	伝記/歴史の注記
3	Subseries#		サブシリーズ番号	17	Creator/ Origination	CREATOR	作成者/出所
4	Box #		箱番号	18	Custodial History	CUST	保管履歴
5	Folder #		フォルダ番号	19	Originals or Copies Note	ORIG	原本または複製の注記
6	Item #		アイテム番号	20	Other Information	ODD	その他の情報
7	Title	Biographical Information	標題	21	Physical Description	PHYS	物理的記述
8	Date		日付	22	Preferred Citation	PREF	望ましい引用形式
9	Description (Scope and contents note)	DESC FIELD	記述 (範囲と内容の注記)	23	Processing Information	PROC INFO	処理情報
10	Access Restrictions	ACCESS FIELD	アクセス制限	24	Related Materials	REL MAT	関連資料
11	Accruals	ACCR FIELD	追加増分	25	UnitID	UNITID	単位識別番号
12	Acquisition Information	ACQ FIELD	入手情報	26	Use Restrictions	USE Restrict	使用制限
13	Alternate Format	ALTFORM FIELD	代替の形式				
14	Appraisal Information	APP FIELD	評価情報				