

京都大学研究資源アーカイブにおける研究資料情報の共有

五島敏芳, 戸田健太郎 (京都大学総合博物館)

Collaborative Compilation and Sharing of Digital Information in Research Resource Archive, Kyoto University Haruyoshi Gotoh, Kentaro Toda (Kyoto University Museum)

In this report, we will talk about Research Resource Archive, Kyoto University (KURRA). KURRA is a project to preserve materials that show educational and research activities in Kyoto University. For this project, we developed a digital archive system called “Peek”.

To archive and digitize materials properly, we need to collaborate with experts in a variety of areas, as there are a variety of activities in Kyoto University. However, the experts are not always in Kyoto University, and collaboration with experts in a distant institution is often not easy. To solve this problem, we developed a workflow for efficient collaboration, and introduced federated identity management (FIM) into Peek.

1. はじめに

1.1. 京都大学研究資源アーカイブ概要

「京都大学研究資源アーカイブ」(Research Resource Archive, Kyoto University : KURRA) は、京都大学における教育研究の過程において作成・収集された資料を保存・活用する、大学全体の活動である(図1)。

対象となる資料は、図書類や標本類と違って、写真・映像・録音、フィールドノート、研究会記録、講義ノート、論文原稿等などの一次資料である。これらの一次資料は、収集・調査を経て、京都大学デジタルアーカイブシステム(Kyoto University Digital Archive System / 愛称 Peek)へ登録され、活用できるように整備される。こうして整備された資料を新たな教育研究の資源(研究資源)と位置づけ、オンラインまたは映像ステーション(Audio-Visual Station : AVS)で公開している。これらの研究資源をもとに京都大学の教育研究活動を紹介する映像・展示等のコンテンツも制作している。オンラインまたは AVS での研究資源公開から、新たに保存・活用の対象となる資料に関する情報が提供されることもある。

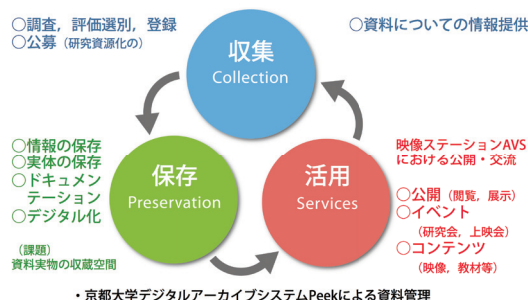


図1 京都大学研究資源アーカイブの活動サイクル

1.2. 教育研究のアーカイブ資料管理の問題点

教育研究に関するアーカイブ資料のばあい、資料管理者(アーキビスト)は、個々の資料の内容を熟知していない／できないことがある。資料整理が外形的把握の水準にとどまることもある。これを補うため、その対象資料の内容を専門とする研究者の協力を得て、資料整理を進める方法がある。京都大学研究資源アーカイブでは、この方法を採用している。

対象資料の専門研究者が、実際に資料を整理する場と近いところに居ればよいが、離れたところに居るとき、資料を整理する際の問題となる。

この問題の解決のため、アーカイブ資料管理手法上の工夫(後掲2章)、情報システム上の環境整備(後掲3章)、を実践した。

2. アーカイブ資料管理手法上の工夫

2.1. 資料調査、整理とデジタル化

アーカイブ資料は、しばしば群として存在する。資料の調査、整理に際して、群として大まかに把握し、現状を写真等で記録する。ふつうアーカイブ資料の現状には、アーカイブ資料のもととなる記録の作成者やアーカイブ資料として伝え残してきた保管者の秩序が残されている。この現状を変更しても元に戻せるようにするためである。

資料の内容をデジタル化する場合、なるべくまとまった分量の資料(群)をデジタル化し、画像や映像等のデータ(デジタルデータ)とする。

これにより、資料に関する情報、記述データ(メタデータ)を採取するとき、1点1点の資料の情報を詳細に採らずに省略することができる。

資料の、現状記録画像、デジタルデータ、メタデータを編集し、資料内容の画像・映像等をとともなう資料目録の構築、具体的には京都大学デジタルアーカイブシステム Peek におけるデジタルコ

レクシオンの構築により、オンライン上での閲覧視聴が可能になる。

2.2. オンライン上のメタデータ作成の協業

デジタル化されたデータを、資料内容に関わる分野の専門研究者に見てもらい、先に省略した資料の詳細な情報、メタデータを作成してもらうことができる。

メタデータ作成は、群のなかのまとまった分量のデジタル化された各資料のすべてを、一人の専門研究者や一つの研究グループだけに、協力依頼しなくてもよい。資料のまとまりのうち、関心のある部分だけでもよく、他の部分は別の専門研究者や研究グループにお願いすることができる。

このとき、メタデータの質は均一である（粒状性が保たれる）必要はない。メタデータがゼロの状態よりも、少しでも作成されることの利益のほうが大きいからである。

3. 情報システム上の環境整備

3.1. オンライン上のメタデータ作成の環境

オンライン上のメタデータ作成の協業（前掲2.2）を実現するためには、2通りの方法がある：a. デジタル化された個々の資料を全世界の誰もが閲覧視聴できるようにする；b. 特定の研究者、研究グループだけに限る。

教育研究資料のなかには、企業秘密的内容や、倫理的にふさわしくない内容、プライバシーに関わる内容などが含まれることがある。

このためbの方法を採用する必要がある。

ふつう web サービスを特定の個人や団体にだけ許可するばあい、専用のアカウント（ID／パスワード）を発行する。

web サービスの提供者にとっては、利用者が多くなれば、アカウント発行したいやアカウント発行のための個人情報管理等派生業務は、かなり煩雑となり大きな負担となる。京都大学研究資源アーカイブのスタッフ規模からは、こうした業務の実現は、むづかしい。

web サービスが多くなれば利用者にとっても、それぞれのサービス毎にアカウントを申請し管理する必要が出てきて、アカウントの管理が負担となる。

web サービスの提供者・利用者双方にとって、各 web サービスで同じアカウントを安全に使用できる、データクリアリングハウスのような仕組みがもとめられる。

3.2. Federated Identity Management 管見

前述の仕組みは、Federated Identity Management（FIM，連合アイデンティティ管理）といい、すでに10年ちかく前から構想が存在していたようだ。日本では2011年頃より商用に登場していたという。いくつかの学問分野でも開発されながら活用が進められてきた。

およそ、現在進行形の研究のデータに対し、関

係各所の研究者複数が持続的にアクセスできるよう、そのために所属機関等既存アカウントを使用できるようにしている。

いくつか事例をあげると、つぎのとおり。
高エネルギー物理学：WLCG（CERN 主導の大型ハドロン加速器の計算処理協業）、生命科学：ELIXIR（汎ヨーロッパ生物学情報研究基盤）、人文科学・社会科学：CLARIN（共通言語資源技術基盤，言語学関係デジタルデータ）、気象学・地球科学：ESGF（地球システムグリッド連合，気候変動データ）。

3.3. 学術認証フェデレーション「学認」の利用

日本におけるFIMの取り組みとして、教育研究分野では2009年より国立情報学研究所と全国の大学等とが連携して構築・運用を開始した学術認証フェデレーション「学認」をあげることができる。電子的学術資源資源を提供または利用する複数個の各教育研究機関（大学，研究所等）が「学認」の規程を信頼しあうことで認証連携を実現し，参加各機関から発行された一つのアカウントで各サービスを利用可能になる，そうした仕組みを「学認」は提供している。

京都大学研究資源アーカイブでは，京都大学デジタルアーカイブシステム Peek に，この「学認」の仕組み，サービスを利用することにした。

これにより，各教育研究機関で公式発行している既存アカウントを利用して，Peek の特定のアーカイブ資料（群）のデジタル化されたデータを，特定の研究者・研究グループだけに提供することが可能となった。

たんに資料を閲覧するだけではなく，メタデータの入力・編集や，デジタルデータの登録も可能である（図2）。ただ，メタデータの項目は固定され（アーカイブ資料管理のメタデータ標準による），入力・編集に際しての操作上の規則に従う必要があり，勝手な入力・編集はできない。



図2 京都大学デジタルアーカイブシステムへログイン後に管理メニューを表示した画面

以上から、京都大学デジタルアーカイブシステム Peek は、一般公開している（ログインなしで閲覧視聴のみ可能な）部分と、「学認」に対応したログインにより使用できる部分とで構成されることとなった。前者を「パブリックモード」、後者を「研究者モード」と呼んでいる。

4. 利益・課題

4.1. 利益

機関公式発行の既存アカウントは、そのアカウント所持者の所属を保証することになる。

「学認」は、各教育研究機関だけでなく、学協会アカウント発行も代行しているという。これにより、「学認」対応のアカウント所持者は、基本的に専門研究者であることが保証される。

これまで教育研究関係のアーカイブ資料の管理において、だれが対象資料の専門研究者か、その範囲を定めづらかった。対象資料の専門研究者は、その対象資料の元記録書類の作成者・保管者と同じ専門分野に限らない可能性があったためである。

「学認」対応のアカウントは、すくなくとも教育研究機関や学協会の所属者で専門研究者と判断できる。かかる専門研究者であることは、取り扱いに注意を要する内容の資料であっても閲覧視聴を許可する条件とできる。

そうした専門研究者により作成されたメタデータは、質的にも信頼性が高い。専門研究者の協力のもとに進めるアーカイブ資料管理には効果が大きい。

4.2. 課題

京都大学研究資源アーカイブでは、「研究資源化プロジェクト」という単位で、1つのコレクションに対し（または1つの出所に収蔵されるいくつかのコレクションに対し）、調査・資料目録作成入力・デジタル化を実施し、京都大学デジタルアーカイブシステム Peek でデジタルコレクションを構築・公開している。研究資源化プロジェクトのメンバーは、京都大学の学内構成員に限られる場合と、学外の関係者も加わる場合とがある。なお、学外者だけで組織される研究資源化プロジェクトは存在しない。

研究資源化プロジェクトが学内構成員だけで組織される場合、京都大学の認証センターの発行した既存アカウントを使って、Peek（研究者モード）へログインし、メタデータやデジタルデータの閲覧、入力編集等を実行できる。実際に、大きな問題は発生していない。

いっぽう学外者をも含む研究資源化プロジェクトでは、その学外者の状況により、いくつかの問題が発生している。

学外者が、「学認」参加機関に所属していない場合、たとえば退職者、名誉教授といった人であれば、認証センター発行の既存アカウントは存在せず、その体系のアカウントとは別に個々のアカウント（またはグループ使用のアカウント）を作成・発行しなければならない。

学外者が、「学認」参加機関に所属している場合でも、その所属機関発行の既存アカウントでログインできないことがある。Peek 研究者モードの要求する情報を、その機関の認証センターが送出許可していないためである。当該の学外者に、所属機関の認証センターへ Peek 研究者モード利用の要望と研究者モード利用に必要な情報の送出の設定の要望を出してもらうしかない。認証関係の情報送出は、その機関の方針・判断しだいだからだ。

京都大学研究資源アーカイブでは、Peek 研究者モードを「学認」対応サービスとするにあたり、「学認」のサイトで送出を要求する情報（属性）を明示している※。それらは、つぎのとおり：organizationName（o，組織名称・英字），jaOrganizationName（jao，組織名称・日本語），eduPersonTargetedID（ePTID，サービス毎の匿名ID）。「学認」参加機関の認証システムが、Peek 研究者モードの要求する情報の送出を許可していれば、ログインの際に送出確認等の画面が現われる（図3）。

※ <https://www.gakunin.jp/participants/> のサービスプロバイダー（SP）の一覧。

図3 京都大学のIdPにおける必要情報送信直前の確認画面

研究資源化プロジェクトに関与していなくても、「学認」参加機関の構成員で、その機関が必要な情報の送出を認めていれば、研究者限定公開のデジタルコレクションを閲覧・検索できる。まだ多くの「学認」参加機関が、Peek 研究者モード利用のための情報送出を認めていない状況にある。各機関で研究者モード利用の要望を出してい

ただければ幸いである。

このほか瑣末な問題として、「学認」を通したログイン過程の画面が煩雑で、詐欺の画面と見誤って警戒される可能性のあることをあげる。

5. おわりに

教育研究世界での FIM は、その事例の多くが研究者にデジタル資源を活用してもらうための環境整備の一部であるように見える。デジタル資源となる／なった研究データは、研究者みずからが所持し、その内容にも熟知している。または研究者から研究データの管理（公開を含む）を託された機関が責任をもって研究データの内容に関するメタデータを整備し、研究者は研究データの管理に煩わされることなく活用できる。要するに、研究者は FIM を通して研究データやそのメタデータを活用するだけである。

教育研究のアーカイブ資料の場合、資料の継承者、所持者と、資料の研究者、資料の内容を熟知する者とが、しばしば一致しない。そのため、研究者による資料のメタデータの遠隔・共同編集が、それを実現する環境とともに必要となった。既存の FIM 事例にみられない試みといえる。とはいえ、ある程度アーカイブ資料の内容をメタデータへ反映しないとデジタルアーカイブシステムにおける資料理解が担保されない、という事例は、アーカイブ資料管理一般にとって参考にならないかもしれない。前述の課題以外にも比較的ちいさな運用上の課題が、まだ存在する（たとえば、所属機関やその部局をこえて複数の研究資源化プロジェクトに関わる研究者の権限設定など）。

こうした京都大学研究資源アーカイブの取り組みは、研究資料情報の共有のあり方としては〈すき間〉の領域を占めるに過ぎないが、現用の研究資料の記録管理が不十分なまま残され非現用へ推移したとき発生する可能性は高く、そうした事態が起これば参考になると予想する。研究資料情報の共有に関心を持つ各所へ協力をあおぎつつ、取り組みを進めていきたい。

引用文献

- Broeder, Daan; Jones, Bob; Kelsey, David; Kershaw, Philip; Lüders, Stefan; Lyall, Andrew; Nyrönen, Tommi; Wartel, Romain; Weyer, Heinz J. Federated Identity Management for Research Collaborations. CERN Preprints, 2012, CERN-OPEN-2012-006, 18p.
<https://cdsweb.cern.ch/record/1442597/files/CERN-OPEN-2012-006.pdf> (参照 2016-01-29).
- 五島敏芳. 京都大学研究資源アーカイブの活動. 記録と史料. 2013 年, No.23, pp.62-64.
- 近藤学. “ID Federation 海外動向/Use Case/トレンド紹介”. Japan Identity & Cloud Summit 2013. 東京, 2013-03-04/05, 国立情報学研究所・OpenID Foundation Japan. JICS (国立情報学研究所), 2013, エンタープライズセッション. <http://id.nii.ac.jp/1125/00000053/> (参照 2016-01-29).
- “概要”. 学術認証フェデレーション. <https://www.gakunin.jp/fed/> (参照 2016-01-29).