
第9章 非文献リポジトリとは

高田 良宏

1. はじめに

国内外に存在する非文献資料（文献以外の学術資料）は膨大かつ多種多様であり、それらに関する資料情報の共有と公開・発信は、学術機関が果たすべき重要な役割の1つである。文献資料においては、金沢大学の KURA をはじめとした全国の学術機関等の機関リポジトリにより情報の蓄積・公開が進んでいる。しかしながら、現在の機関リポジトリは、文献を取り扱うために発展してきたものであり、非文献資料情報を蓄積し、公開するには不十分である。そのため非文献資料の蓄積・公開に適した非文献リポジトリの確立が望まれる。今回は、先行する機関リポジトリの例(KURA 等)と、発展途上であるが我々が研究開発を進めている非文献リポジトリを取り上げ解説する。

2. 機関リポジトリ

機関リポジトリとは、大学等の学術機関の構成員が生産した様々な学術情報を電子的形態で収集・蓄積し、インターネット上で公開するためのシステム（サービス）である。機関リポジトリは、大学図書館が中心となって構築される場合が多いが、その経緯は学術コミュニケーションの危機的状態（雑誌の危機：シリアルズ・クライシス）と大学からの情報発信の強化にあるといわれる[1]。国内では国立情報学研究所（NII）が中心となり、学術機関リポジトリ構築連携支援事業として推進してきた[2]。当事業の Web サイトの情報によれば、国内の 216 もの機関（共同体も含む）が機関リポジトリを運用している（2013 年 1 月 26 日現在）。

機関リポジトリの対象は、学術機関の構成員が生産した様々な学術情報とされているが、現実には学術論文、紀要、研究報告書などの書誌系の情報が中心である。書誌系の情報以外の写真、動画、観測データをはじめとした多くの貴重な学術情報は対象外とされる場合が多い。3 節でも説明するが、書誌系の資料を文献資料（文献コンテンツ）、書誌系の資料以外を非文献資料（非文献コンテンツ）と呼ぶ[3]。

一般的な機関リポジトリの例として、金沢大学学術情報リポジトリ（KURA :

Kanazawa University Repository for Academic Resources) [4]を、さらに KURA をはじめとする日本の機関リポジトリに蓄積された学術情報を横断的に検索可能なリポジトリポータル JAIRO (Japanese Institutional Repositories Online) [5]を紹介する。

● 金沢大学学術情報リポジトリ (KURA)

KURA は 2006 年に、金沢大学における教育・研究活動から生み出された学術的な成果を電子的な形態で保存・公開するシステムとして稼働を開始した (図 1)。以下、KURA の概要を説明する。なお、他機関の機関リポジトリも、その目的や対象、および、操作方法等においては、KURA と大きな違いはないと考えてよい。



図 1 KURA のトップページ

KURA の対象：

- ①学術雑誌掲載論文：本学教職員等が、学外の学術雑誌等に発表した論文や報告
- ②紀要：本学刊行の紀要・広報等
- ③報告書：本学刊行の COE 等の報告書。今後、科研費報告書も収録予定
- ④学位論文：本学に提出された学位論文で著者の許諾の得られたもの（今後義務化される予定）
- ⑤その他：会議・学会等の発表資料等。その他、プレプリント、実験データ、電子教材、シラバス、統計データなどについても可能

KURA に登録された学術情報の検索方法：

KURA に登録された学術情報資料を検索するには、キーワードで検索する方法と著者の所属等の階層をたどる 2 通りの方法がある (図 2)。

検索結果一覧(キーワードでの検索)

発行日	タイトル	著者
2005年 3月 31日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
1993年 3月	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
1989年	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
2004年 3月 5日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
1991年 3月 28日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide

検索結果一覧(階層をたどっての検索)

発行日	タイトル	著者
2011年 4月 28日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
2012年 1月 15日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
2011年 4月 30日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide
2011年 4月 28日	「ワジュラヴァナー」(彩色の巻)和訳	森 雅彦; Mori, Masahide

図2 学術情報資料の検索法
(上段:キーワードで検索する方法, 下段:所属等の階層をたどって検索する方法)

アイテム(登録されている資料情報)表示画面:

図3にアイテムの表示画面(簡易表示版)を示す。資料に対する情報(メタデータ)の記述は、Dublin Core[6](以下DC)に準拠している。そのため世界中の機関リポジトリとデータの互換性があり、後述のハーバスティング[7]により、他リポジトリとの連携が可能である。また、ハンドルという世界中で一意的識別子が割り当てられている。ハンドルはURL形式になっており、資料を保存しているサーバが移転しても、ハンドルが研究成果を識別するので、永続的なアクセスが可能となる。

全文を表示する

ファイル 記述 サイズ フォーマット

g02.pdf 1.12 MB Adobe PDF

Dublin Core

タイトル: 汎用データフォーマットを利用した自然科学データアーカイブシステムの開発

著者: 尾崎, 友紀
高田, 良宏
立原, 清也

発行日: 2006年

出版社(者): 学術情報地理研究編集委員会

雑誌名: 学術情報地理研究 = Journal for Academic Computing and Networking

ISSN: 1343-2915

巻: 10

号: 1

開始ページ: 5

終了ページ: 14

URI: http://hdl.handle.net/2297/7030

資料種別: Journal Article

版表示: author

世界中で一意的識別子

図3 アイテムの簡易表示画面

● JAIRO

JAIRO は NII が運営する機関リポジトリポータル（機関リポジトリのポータルサイト）で、国内の機関リポジトリに蓄積された学術情報を横断的に検索することができる（図 4）。KURA に登録されている情報も JAIRO で検索することができる。JAIRO は、ハーベスティングにより、各機関リポジトリから DC で記述されたメタデータを収集・蓄積し、公開している。図 5 にハーベスティングの概念を示す。各機関リポジトリは登録されている資料に対する情報（メタデータ）を生成し、JAIRO の要求に回答して提供する。JAIRO には各機関のメタデータのみが収集されている。資料はハンドルにより各機関リポジトリで参照することが可能である。



図 4 JAIRO のトップページ

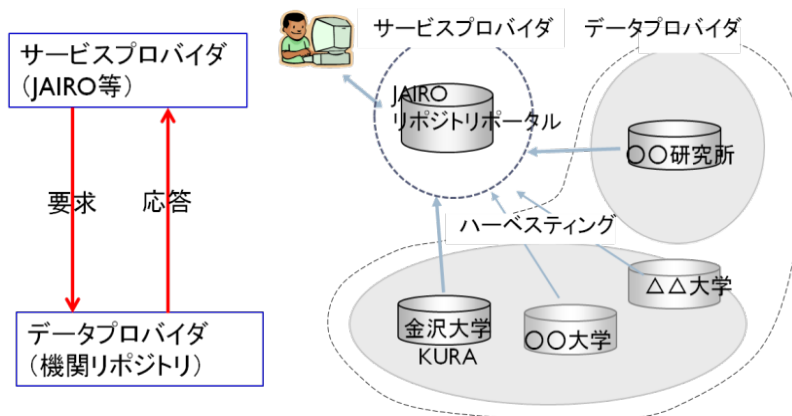


図 5 ハーベスティングの概念

3. 非文献リポジトリ

非文献リポジトリとは、非文献資料を対象としたリポジトリで、非文献資料の取り扱いが管理者にとっても利用者にとっても容易なものでなければならない。2節でも述べたが、大学内に蓄積している非文献資料の多くは、その所有者や利用者、さらに、機関リポジトリを運用する図書館がリポジトリ化を強く望んでいるにも関わらず、リポジトリ化の対象外とされる場合が多い。第1の理由は、資料の種類が多様なため、主に書誌情報を元に情報管理を行うリポジトリへの登録にうまく適合しないことが挙げられる。また、これらの資料の多くは、文字情報を含まず、全文検索が利用できないため、付加されたメタデータに対する検索が中心となり、検索性が低いと言わざるを得ない。このため、資料登録者には適切なメタデータ付けが、利用者には適切な検索語句の入力が求められ、これもリポジトリへの登録を阻む原因となっている。そのため非文献リポジトリの構築・運用方法の確立が望まれているわけだが、非文献リポジトリが対象とする資料には、当然文献資料も含まれる。非文献リポジトリという言葉は、非文献リポジトリの構築・運用方法が確立され、それが学術機関に普及すれば消滅するだろう。

現状で、非文献リポジトリ構築し、運用する場合、次のような問題点を解決する必要がある。

①プラットフォームの課題

多種多様な非文献資料の取り扱いに適した汎用プラットフォームが必要

→ 機関リポジトリで多用されているプラットフォームは非文献資料の取り扱いを考慮していない

②メタデータの課題

非文献資料の多種多様かつ専門的な情報を的確に表現するためのメタデータ要素群の設計が必要

→ 機関リポジトリで一般的に使われているメタデータ要素群は非文献資料の取り扱いを考慮していない

③横断検索の課題

分野、機関横断的検索環境が必要

→ 機関リポジトリポータルである JAIRO の非文献資料版が必要

我々は、非文献リポジトリの構築・運用方法の確立を目指し、2007年に研究活動を開始した。途中活動形態に変化があったものの現在も活動は継続中である。これまでの取り組みを4節で、現在の取り組みを5節で紹介する。

4. 金沢大学での取り組み（これまでの取り組み）

非文献資料をリポジトリ化する場合の指針となるべく、KURA と連携可能な機関リポジトリプラットフォームの一つである DSpace[8]を改良し、可視性と保守性に優れた汎用性の高い学術情報リポジトリの構築を目指した[3]。その概要を 4.1 節から 4.7 節で述べる。

4.1 開発方針

表 1 に開発要件を示す。なお、本開発で解決すべき課題として挙げた項目を表中に課題 1～課題 4 として示す。

表 1 開発要件

<p>①汎用性の確保</p> <ul style="list-style-type: none"> ●メタデータの互換性が確保できること ⇒ 当該リポジトリ上での詳細な定義と、他リポジトリとの互換性を両立（課題 1） ●他リポジトリとの連携を行えること ⇒ OAI-PMH プロトコルを活用した他リポジトリとの連携を実現（課題 4） <p>②保守性の確保</p> <ul style="list-style-type: none"> ●複数の異種資料の管理を容易に行えること ⇒ 同一リポジトリに異種資料を容易に共存させる環境を導入（課題 2） ●多様かつ膨大な数の資料の管理を容易に行えること ⇒ 分類の登録・管理機能、一括登録機能を導入（課題 2） ●プラットフォームの保守性が高いこと ⇒ 既存リポジトリプラットフォームをベースとする <p>③可視性の向上</p> <ul style="list-style-type: none"> ●資料が持つ発掘地、所蔵地などの地理的位置に関する情報を可視化し、視覚的な検索機能を提供すること ⇒ 地理的位置情報と Google Earth を連携させた検索機能を導入（課題 3）

4.2 対象資料

アジア図像集成[9]（図 6 左上）として蓄積しているインドの仏像・壁画・遺跡に関する画像を対象にして開発および検討を行った。また、本システム開発後の検証の目的で、e-Learning 素材（図 6 左下）、四高物理機器（図 6 右上）および「あけぼの衛星」の観測データ（図 6 右下）に適用した。

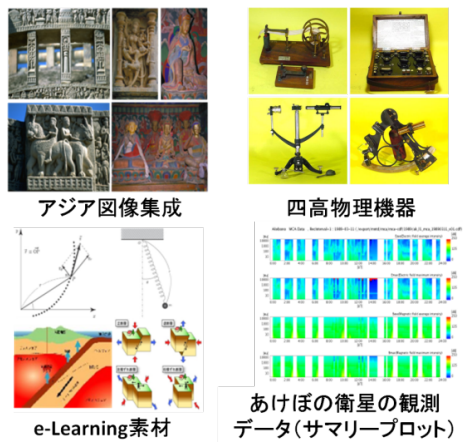


図6 対象とした資料の例

4.3 システム概要

本システムは、リポジトリプラットフォームのDSpaceをベースとして、機能を改良、追加する形で開発を進めた。開発にあたっては、DSpaceの既存クラスを極力書き換えず、メソッドのオーバーライドおよび外部スクリプトを追加することとし、システムの移植やDSpaceのバージョンアップといった場合にも極力影響が出ないように配慮した。図7に、実装したアジア画像集成の概要(画面遷移の例)を示す。



図7 実装例(アジア画像集成)

4.4 メタデータの互換性の確保

前述の通り機関リポジトリではDCと呼ばれるメタデータ語彙を使用している。DCは15の基本要素と、より詳細な情報を表すための限定子から構成されている。本来DCは書誌情報を記述するためのメタデータであるため、そのままでは非文献資料に適さない。そのため、前述の資料に対して、互換性を保ちつつ必要な項目を追加して拡張メタデータ語彙を定義した。すなわち、限定子を取り除いた場合でも基本要素と値との間に矛盾がないというDumb-Down原則に従って定義することであり、これによりリポジトリ上での詳細な定義と、他リポジトリとの互換性を両立させることができる。

4.5 保守性の確保

異種資料の共存

異種資料を同一リポジトリ上に共存させるには、資料ごとに適したメタデータ語彙を設定できることと、資料ごとに適した一覧表示を行える必要がある。メタデータ語彙の設定に関しては、登録するそれぞれの資料に適したメタデータ語彙がその種類分設定可能で、さらに、登録の際に容易に切り替えて利用できるようにした。一覧表示に関しても、それぞれの資料に適したメタデータ項目の設定が資料の種類分設定可能とし、表示の際に利用できるようにした。

コミュニティとコレクションの管理

DSpace はコミュニティやコレクションで資料の実態（アイテム）を分類し、資料の管理を行っている。これらの作成は Web インターフェイスに頼っており、資料の登録とともに負担となる部分である。ここでは、Excel ファイルを用いてコミュニティとコレクションの構造の管理を行えるようにした。この仕組みを導入することで、DSpace の管理者のコミュニティとコレクション管理に掛かる負担が軽減されるとともに、リポジトリの構造の再現を容易に行うことが可能となった。

一括登録

非文献資料の場合、アイテムが数千件、数万件以上におよぶ場合が多く、これらの登録を簡単な準備で一括登録できるスクリプトを開発した。Excel などの表計算ソフトの出力ファイルを読み込み、一括登録が行える仕組みを導入した。具体的には、メタデータ、アイテムおよび登録先のコミュニティとコレクションなどの情報を記述した TAB 区切りまたは CSV 形式のファイルの内容を読み込み、メタデータなどの解析を行い、DSpace が理解できる形式に変更し登録する。この方法を導入することにより、登録の際の負担が軽減されるとともに、資料ごとの再現を容易に行うことが可能となった。

4.6 Google Earth による情報の可視化

文字情報を持たない非文献資料では、全文検索が利用できないため、資料に付与されたメタデータのみによる検索が行なわれる。そのため利用者は文献資料の検索と比べると、より適切な検索語句を指定する必要がある。一方、今回利用したアジア画像集成は、画像の出土地や所蔵地といった地理的な位置（地名）を示す情報を持っている。また、アジア画像集成に限らず、学内に存在する非文献資料の中には、地理的な位置情報を持つ資料が少なくない。これら資料が持つ地理的な位置情報を利用し、

Google Earth と連携して、位置情報を地図上に表示することで可視化し、資料の検索性の改善を図った。今回、Google Earth との連携には KML (Keyhole Markup Language) を用いた。Google Earth 上に表示させる情報には、地名とアジア画像集成へのリンクを含めることで、Google Earth 上からアジア画像集成への検索も可能となった。図 8 に Google Earth と連携した可視化システムの概要、図 9 Google Earth での表示例を示す。

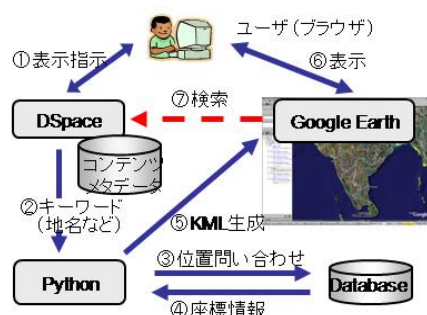


図 8 Google Earth と連携した可視化の概要



図 9 Google Earth 上での表示例

4.7 他のリポジトリとの連携

学内の学術情報を統一的に公開するリポジトリポータルを試作を行なった。図 10 にその仕組み（ハーベスティングにより統合検索の概念）を示す。これは、OAI-PMH プロトコルに従い、学内のリポジトリのメタデータをハーベスタにより収集するものである。今回は非文献資料に最適化した当システムと文献資料に適した図書館の KURA で実証し、文献・非文献資料を統一的に検索することができた。これにより、当システムと現況のリポジトリ（KURA など）が連携可能であることを示すことができた。

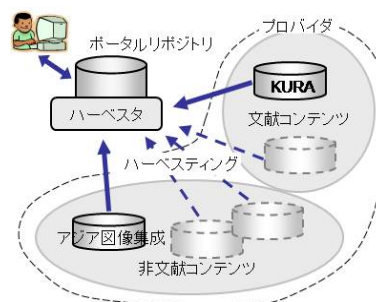


図 10 ハーベスティングによる統合検索

5. 金沢大学での取り組み（現在の取り組み）

4 節の取り組みで、非文献リポジトリ構築に関する指針を示すことができた。また、資料を取り扱う関係者に対しては非文献リポジトリの重要性・有用性を認識させることができたと考える。この取り組みは、現在、金沢大学資料館 Virtual Museum Project[10]

（以下、VMプロジェクト）と学術資源リポジトリ協議会[11]に引き継がれている。VMプロジェクトは、資料の「アーカイブ化」と「展示公開・コンテンツ化」という2つの対極にあるコンセプトの融合を目指している。現在、表2に示す資料がリポジトリ化されている。図11にその表示例を示す。これらは資料館のVMプロジェクトのWebサイトで閲覧可能である。リポジトリプラットフォームに関しては、我々も共同研究という形で開発に参加しているNIIのWEKO[12]を使用している。日本発の非文献資料に適したリポジトリプラットフォームとして期待したい。また、メタデータに関してはKUVMモデルを考案し学会等で議論を続けている。

学術資源リポジトリ協議会は、広く非文献学術資料を対象にした横断的な情報共有基盤の構築・整備と、それらに関わる人的なネットワークの形成を目的とする独立した協議会組織である。非文献資料や非文献リポジトリに関心のある研究者、博物館・資料館の実務者、関連ソフトウェア開発に携わるNII・企業・NPO等の実務者などによって構成されている。

表2 リポジトリ化された資料

①石川師範学校写真資料
旧制石川師範学校に伝来した古写真。大正・昭和初期の修学旅行の様子等を撮影している。金沢大学資料館蔵
②金沢大学医学図書館図面
昭和44年(1969)11月竣工の金沢大学医学図書館設計図控。一部同じ図面のため画像未掲載。金沢大学附属図書館蔵
③金沢病院設計図
明治38年(1905)8月、金沢市小立野に完成した石川県立金沢病院(後金沢大学附属病院)設計図。金沢大学資料館蔵
④きのこムラージュ標本
ムラージュは鋳型を用いた蠟模型のこと。当標本は植物学や博物学の教育用標本として作製。金沢大学資料館蔵
⑤第四高等学校物理実験機器
旧制第四高等学校で高等教育に使用されていた物理実験機器。金沢大学資料館蔵
⑥第四高等学校教育掛図
旧制第四高等学校で高等教育に使用されていた教育掛図。金沢大学附属図書館蔵。現在は植物・地史110点のみ公開

6. 他大学等での例

他大学等での例を付録の5に示すので参照されたい。



図 11 VM プロジェクト

参考文献

- [1] 尾城孝一: 「学術機関リポジトリ」, 変わりゆく大学図書館, 逸村 裕, 竹内 比呂也編, 勁草書房, pp.101-114(2005)
- [2] 国立情報学研究所(NII): 「学術機関リポジトリ構築連携支援事業」, <http://www.nii.ac.jp/irp/>
- [3] 高田良宏, 笠原慎也, 西澤滋人, 森雅秀, 内島秀樹, 非文献コンテンツのための可視性と保守性に優れた学術情報リポジトリの構築, 情報知識学会誌, 19(3), 251-263, 2009-10-28.
- [4] KURA, 金沢大学, <http://dspace.lib.kanazawa-u.ac.jp/dspace/>
- [5] JAIRO, 国立情報学研究所(NII), <http://jairo.nii.ac.jp/>
- [6] Dublin Core Metadata Initiative, <http://www.dublincore.org/>
- [7] Open Archives Initiative: “Open Archives Initiative Protocol for Metadata Harvesting”, <http://www.openarchives.org/pmh/>
- [8] DSpace Foundation: “DSpace.org”, <http://www.dspace.org/>
- [9] 金沢大学: 「アジア画像集成(Asian Iconographic Resources)」(DSpace 版), <http://wwdb02.db.kanazawa-u.ac.jp/dspace/>
- [10] 金沢大学資料館 Virtual Museum Project, <http://kuvm.kanazawa-u.ac.jp/>
- [11] 学術資源リポジトリ協議会, <http://srv1.amane-project.jp/hibunken/htdocs/>
- [12] WEKO, <http://weko.at.nii.ac.jp/>