

ジャパンサーチが目指すもの

——国立情報学研究所コンテンツ科学研究系教授 高野明彦氏による発表の報告

山本 睦月(立命館大学大学院文学研究科)

E-mail <mailto:gr0517fi@ed.ritsumei.ac.jp>

1. 概要

本稿は、2021年10月6日に行われた「第91回国際ARCセミナー」における高野明彦氏(国立情報学研究所コンテンツ科学研究系教授)の発表について報告するものである。本発表では、発表者が開発当初から関わる「ジャパンサーチ」¹⁾の現状と機能、今後目指していくミッションについて述べられた。

2. ジャパンサーチとは何か

2-1. ジャパンサーチの概要

ジャパンサーチとは、国立国会図書館がシステム運用・連携協力を行うデジタルアーカイブの検索・閲覧・活用基盤であり、様々な分野のデジタルアーカイブと連携し、我が国の多様なコンテンツとそのメタデータをまとめて検索・閲覧・活用することを可能にしている。ジャパンサーチでは研究者と各アーカイブ機関との連携をサポートし、一人の研究者が同時に複数のアーカイブ機関の所蔵品を活用することや、複数の研究者同士の共同研究のプラットフォームになることで、様々な研究者の手助けしている。

政府の「知的財産推進計画」²⁾に掲げられている国の取り組みに支えられ、2019年2月に試験版が運用開始、2020年8月に108件のDBと約2100万件のメタデータを持つ状態で正式版として運用が開始された。

2-2. ジャパンサーチの連携方針

ジャパンサーチにデータを入力するには2つの方法がある。

まず、既に確立されたアーカイブ機関であれば直接連携をすることが可能だ。国の機関など当該分野におけるコンテンツを幅広くカバーしている機関に該当する。

次に、上記に該当しない機関は「つなぎ役」を立てて連携を行う。これがジャパンサーチの連携方針としては原則となっている。つなぎ役はメタデータの取りまとめ・標準化・共有化を図り、コンテンツへの長期アクセスを保証する基盤としての役割を担う³⁾。この方法があることで、各所蔵館や分野を超えた統一性を持つ理想

的なメタデータを追求することが可能になると発表者は語る。システムは各所蔵館が個別に抱える問題と密接に関わるために幾つかのパターンが準備されており、フィードバックによって随時変更される⁴⁾。現在はつなぎ役としての役割を持つ連携機関は29館あり、国、地方自治体、大学、公益法人・民間機関などが参加している⁵⁾。

2-3. メタデータと共通項目

現在、ジャパンサーチには141件のデータベースと約2340万件ものメタデータが登録されている⁶⁾。多岐に渡る分野のデータが収集されていることこそがジャパンサーチの価値の一つであると発表者は指摘する。

また、ジャパンサーチではメタデータとサムネイル(プレビュー)、コンテンツの3つの情報資源をそれぞれ重要視しており、メタデータとサムネイルをオープンに流通させることでコンテンツの利用が促進されるとも指摘する。ジャパンサーチの場合ではサムネイルやコンテンツは各機関が公開している先のURLをメタデータの一部として収集し利用しているため、各機関にはメタデータはもちろんサムネイルの公開を特に要請している。

こうして収集したコンテンツには権利区分を明確にすることも各機関に要請し利用者がどこまで利用可能なのかを掲示している⁷⁾。

収集したメタデータには、「共通項目ラベル」⁸⁾を付与し、分野横断の串刺し検索を実現している。これは従来の単一フォーマットにマッピングする方法ではなく、個別のデータの持つ価値はそのままに分野横断検索をするときに必要となる共通項目を新たに付与し、検索に用いる方式だ⁹⁾。

3. ジャパンサーチの機能

上記のようにして収集されたデータには、大きく分けて下記の3つの観点から機能が追加されている。

3-1. 検索機能

まず重要な機能として検索機能が挙げられる。ジャパンサーチでは全データベースをキーワード検索の他、上記の共通項目ラベルを活用した項目別検索も可能であり、横断検索・詳細検索に力を入れている。

また、検索対象とするデータベースやメタデータの項目を定義し検索制約をかけて行うテーマ別検索は、検索する範囲を狭めながらもテーマに沿った多様な資料を検索できる機能だ¹⁰⁾。

さらに AI を用いた画像検索では、サムネイルなどの画像から類似検索が行えるようになっている。

3-2. 楽しめる機能

ジャパンサーチには検索だけでなく「楽しむ」機能も追加されている。その一つが「ギャラリー」機能だ。所謂デジタルミュージアムに値するものもあり、現在は約300テーマが公開されている¹¹⁾。検索機能を利用しない場合でも、見ることで楽しめる、また検索のヒントにもなるページだ。

ここでは IIIF の機能が非常に有効的に活用されている。IIIF とは国際的な画像の相互運用のための規格であり、画像を発信する側と受信する側の双方が同じ規則に従う¹²⁾。ギャラリー機能もこの規格に即して運用されている。

ギャラリー機能は他の利用者が検索した結果を共有するものであり、手動でやることの面白さや深さを大切にしていこう場所としての機能だと発表者は語る。

3-3. 利活用機能

検索機能と楽しむ機能を活用し、新たなアクティビティのプラットフォームを作る機能もジャパンサーチには備わっている。

「マイノート」機能は個人で検索した履歴を残しそれを共有できる機能だ。メタデータやギャラリー、検索結果を登録するだけでなく、注釈をつけることも可能でありメモ機能としての役割も果たす。そしてその結果を様々な形式¹³⁾でエクスポートすることが可能であり、他の利用者と共有することができる。

これを複数人で同時に編集できる機能が「ワークスペース」機能だ。連携機関のみが作成可能な機能だが URL とパスワードを知っていれば誰でもアクセスが可能であり、成果物はジャパンサーチ上でギャラリーとして公開できる他、ウェブパーツとして外部に貼付することも可能だ。これらは調べ学習やキュレーション実習などでの活用が視野に入れられている機能である。

また、任意の研究者たちによって組織される「プロジェクト」は連携機関ができる一通りの作業ができる機能として人気を高めている機能だ¹⁴⁾。

4. ジャパンサーチが目指すもの

さて、先日発表者らは「ジャパンサーチ戦略方針2021-2025」¹⁵⁾を打ち出した。これはジャパンサーチの今後だけでなく、ジャパンサーチがこの社会に起こした変化そのものを考え、そのためにジャパンサーチをどうしていくかの方針を定めるものであった。キャッチフレーズは「デジタルアーカイブを日常にする」¹⁶⁾。

方針内で掲げられたミッションは以下の通りだ。

新しい情報技術とアーカイブ連携を通じて、日本の文化的・学術的コンテンツの発見可能性を高め、それらを活用しやすい基盤を提供することで、デジタルアーカイブが日常に溶け込んだ豊かな創造的社會を実現します¹⁷⁾

これを主軸に、ジャパンサーチを利用した活動の柱が決まった。「支える」「伝える」「広げる」「挑む」の4つのアクションだ¹⁸⁾。これらはデジタルアーカイブの持つ本質的な価値を最大化するために設定されたものである。

では価値とは何か。ここで唱えられる価値は「記録・記憶の継承と再構築」「コミュニティを支える共通知識基盤」「新たな社会ネットワークの形成」の3つだ¹⁹⁾。自身の研究範囲など限定されたコミュニティによって共有されている知識基盤ではなく、広がりを持った知識基盤を作ること、そして隣接するコミュニティに気づくことが重要だと発表者は指摘する。分断された知識基盤ではなく、より柔軟な広がりのある知識基盤によって新たなコミュニティを形成していくことは、自身の研究の便利さの追求以上にデジタルアーカイブが持つ本質的な価値であり、それを行う場所としてジャパンサーチを活用してほしいと発表者は語る。自身の専門を一度忘れ、自身が抱いた問題意識を掘り下げる場所としてジャパンサーチを利活用し新たな知見と出会うことは、自身の専門家としての枠組みを抜けることに繋がるからだ。

[注]

- 1) ジャパンサーチ「ジャパンサーチ」
<https://jpsearch.go.jp> (閲覧日:2022年1月10日)
- 2) 内閣府 知的財産戦略推進事務局による発表。詳細は以下冊子を参照されたい。
内閣府 知的財産戦略推進事務局 「知的財産推進計画2017(2017年5月16日知財戦略本部会合決定)」概要
https://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikeikaku20170516_gaiyou.pdf (閲覧日:2022年1月10日)
内閣府 知的財産戦略本部 「知的財産推進計画2017」
<https://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikeikaku20170516.pdf> (閲覧日:2022年1月10日)

- 3) ヨーロッパで文化遺産資料のデジタルアーカイブを推進しその基盤となっている「Europeana」では、「aggregators」と呼ばれるつなぎ役を介して各アーカイブ機関からデータを収集し活用者にとって価値のあるものを作っている。ジャパンサーチでもこの手法が取り入れられた。
Europeana 「Europeana Aggregators Forum」
<https://pro.europeana.eu/page/aggregators> (閲覧日:2022年1月10日)
- 4) 現在は以下の4つのパターンが準備されている。「1,同分野・地域内のアーカイブ機関のデータベースをまとめるポータルを構築」「2,同じ組織内の複数機関のデータを集約したデータベースを構築」「3,同じ分野・テーマの資料を様々な団体又は個人から集約したデータベースを構築」「4,データベースをもたないが、つなぎ役としてアーカイブ機関への情報提供、事務手続き、メタデータ標準化等をサポート」
- 5) 2021年9月21日時点。立命館大学アート・リサーチセンターは私立大学の一つとして参加している。
- 6) 2021年9月21日時点。
- 7) 「教育利用」「非商用利用」「商用利用」の大きく3つに分野を分け、それぞれ権利区分を指定するように要求している。また、これらの権利区分は「トップ画面」「検索結果での絞り込み」「個別コンテンツ(早見表)」などで掲示している。
- 8) 名称・タイトル、時間、場所、所蔵機関、URL 等の分野横断で共通となりうる項目のみ。この共通項目ラベルと収集したメタデータを「ジャパンサーチ利活用スキーマ」と呼ばれる、ジャパンサーチにおけるメタデータの標準形式に変換する。これは Europeana 等とも連携が可能だ。
- 9) 収集したメタデータにある個別項目は検索先で閲覧可能にしている。
- 10) 検索制約(フィルター)は編集画面で作成する。連携機関のみ作成可能だ。
- 11) 2021年9月21日時点。立命館大学アート・リサーチセンターは「日本の伝説 異界」というギャラリーを公開している。
立命館大学アート・リサーチセンター・国立国会図書館作成「日本の伝説 異界」
<https://jpsearch.go.jp/gallery/ndl-9DnAAqJpV6>
(閲覧日:2022年1月10日)
- 12) IIIFの詳細については、以下のウェブサイトを参照されたい。
International Image Interoperability Framework 「How It Works」
<https://iiif.io/get-started/how-iiif-works/> (閲覧日:2022年1月10日)
- 13) データを CSV、Excel、JSON 形式でエクスポート可能。ウェブパーツとして外部サイトに貼付も可能。
- 14) プロジェクト上では、データベースの公開、ギャラリーやワークスペースページの作成等の作業が可能であり、プロジェクトの成果物はジャパンサーチ上で一般公開することも、限られたなかで共有することも可能である。しかし、プロジェクト上のデータはジャパンサーチの横断検索や一覧表示の対象外となっており、研究プロジェクト等でデータベースの登録やギャラリーの試作の場としての利用

が想定された機能だ。

- 15) ジャパンサーチ「ジャパンサーチ戦略方針 2021-2025 「デジタルアーカイブを日常にする」」
<https://jpsearch.go.jp/about/strategy2021-2025>
(閲覧日:2022年1月10日)
- 16) ジャパンサーチ、同上
- 17) ジャパンサーチ、同上「ミッション」
- 18) ジャパンサーチ、同上「4つのアクション:ジャパンサーチを使った活動の柱」
- 19) ジャパンサーチ、同上「3つの価値:デジタルアーカイブの大切な役割」

[参考文献]

立命館大学アート・リサーチセンター「第91回 国際ARCセミナー (Web 配信)」
<https://www.arc.ritsumei.ac.jp/j/news/pc/009272.html> (閲覧日:2022年1月10日)