

〈ARC のデータベースの基本解説 その2〉

◇ARC ポータルデータベースから IIIF マニフェストを取り出してみよう◇

1. IIIF の広まり
2. 相互運用性と研究用データベースのジレンマ
3. IIIF マニフェストへの自動マッピング API

1. IIIF の広まり

メタデータの標準化や相互運用性が叫ばれるようになって久しくなりました。人文学研究の世界でも、メタデータ規則として Dublin Core が周知されるようになってから、およそ 20 年が経過しようとしています。近年では、IIIF (International Image Interoperability Framework) が注目を浴び、デジタル画像を中心とした国際的な相互運用の枠組みが実際に運用されるようになり、さらに広まりを見せています。

IIIF では、コンテンツの全般的な情報(メタデータや画像の情報)を記述した IIIF マニフェストと呼ばれる設定ファイルを、Mirador、Curation Viewer といった IIIF ビューアで読み取ってコンテンツを閲覧します。ユーザー視点では、こうした取り組みによって、各所蔵機関が独自に運用するデータベースの検索方法や画像閲覧方法のクセを学習する必要がなくなり、画一化された方式で資料の閲覧ができるようになるメリットの大きさは計りしれません。

2. 相互運用性と研究用データベースのジレンマ

立命館大学アート・リサーチセンター(ARC)では、1998年の設立以来20年にわたって、ARCを含む国内外の諸機関が所蔵する文化資源をデジタル化し、研究素材として蓄積し、それらのデータベースを構築してきました。もちろん ARC 所蔵資料については、データベースが所蔵品目録の役割を果たし、一般の外部ユーザーに所蔵品の情報を伝える役割を持っているわけですが、ARCは設立当初から、研究機関として研究者用のデータベースを構築することを基軸として、データベース開発と蓄積を進めてきました。つまり、一般向けの分かりやすさは追求せず、研究者視点のメタデータ構成と、研究ツールと呼び得るデータベース機能を持ったデータベース開発を行ってきた経緯があります。

研究者は、本来的にその専門分野のマニアですから、データベース開発において研究者視点を強くすれば強くするほど、メタデータにはマニアックな視点を取り込まれ、分野の垣根を乗り越えていくような標準化や相互運用性から必然的に遠ざかっていきます。もちろん、単一のデータベース上で研究が完結するならば、それでも良いかもしれません。しかし、デジタル化された文化資源を各機関がオンライン公開し、関連分野のデジタル文化資源をも包摂しながら研究活動を行う必要性や、それらをオンライン上でキュレーションすることが必須となってきた現状では、標準化や相互運用性は担保しておかなければならない課題でしょう。しかしまた一方では、IIIF の思想や機能を一から学び、それらを意識しつつ、メタデータ構成やデータベースを再構築しながら研究を進めるのは、文系研究者にとって大きな負担であり、現実的には難しいと考えられます。

3. IIIF マニフェストへの自動マッピング API

上述の観点から、テクニカルサポートボードでは、ARC のデータベースに蓄積したデータを、自動的に IIIF マニフェストにマッピングする API(以下、変換 API)を開発し、実装を開始しました。ARC のポータルデータベースには、個々のレコードに permalink(パーマリンク、個々のレコードへの固定リンク)が設定されています。例えば、浮世絵ポータルデータベースでは、詳細情報表示画面の下部に

<https://www.dh-jac.net/db/nishikie/arcUP0001/portal/>

とあるのが permalink です。この permalink の末尾に、文字列「manifest.json」を付加し、

<https://www.dh-jac.net/db/nishikie/arcUP0001/portal/manifest.json>

とすれば、それがそのまま IIIF マニフェストになります。古典籍ポータルデータベースの場合は、各冊の詳細書誌画面の下部に

<https://www.dh-jac.net/db1/books/arcBK01-0001/portal/>

とあるのが permalink です。この末尾に「manifest.json」を付加し、

<https://www.dh-jac.net/db1/books/arcBK01-0001/portal/manifest.json>

とすれば、IIIF マニフェストになります。この IIIF マニフェストをウェブブラウザで開くと、標準化されたメタデータと画像の格納先などの情報が、JSON 形式という機械判読可能な形で記述されていることがお分かりいただけるかと思います。この IIIF マニフェスト(マニフェスト URL)を上記の IIIF ビューア上で指定すれば、ARC のウェブデータベース上でなくとも、該当の資料を閲覧することが可能になり、相互運用性を担保することが可能です。

ARC は、ポータルデータベースに収録される一般公開データについては、国立国会図書館が運用している国の分野横断型統合ポータル「ジャパンサーチ」(<https://jpsearch.go.jp/>)との連携を進めていますが、ジャパンサーチにも IIIF マニフェストを提供しているため、ジャパンサーチ上で ARC のポータルデータベースに収録される資料を閲覧できるようになっています。

閲覧だけではありません。ARC の諸行事でお知らせしてきましたように、現在、ARC では、オンラインで稼働する研究ツール「Kinukake」を鋭意開発しています(2021 年度中の一般公開予定)。今後、本誌に詳細な報告が掲載されると思いますが、Kinukake は、オンライン上にあるデジタル文化資源を読み込み、個々の資源に解説やアノテーションを付したり、資源をグループ化したり、相互に関連づけるなどしながら、知的生産活動を行うことができるリサーチスペースとして構築を目指しています。この Kinukake へのデジタル文化資源の読み込みも、IIIF マニフェストを介して行います。IIIF マニフェストによって、研究に関わる素材を Kinukake のような研究ツール上に収集し、その上で分析・考察するような、ポーンデジタル型研究環境が整いつつあるのです。

本稿執筆時点において、変換 API 実装済のデータベースは下記の通りです。

- ARC 浮世絵ポータルデータベース
https://www.dh-jac.net/db/nishikie/search_portal.php
- ARC 古典籍ポータルデータベース
https://www.dh-jac.net/db1/books/search_portal.php
- ARC 近代書籍ポータルデータベース
https://www.dh-jac.net/db1/mbooks/search_portal.php

その他の基幹ポータルデータベースや、ポータル以下の入口(専門別入口)についても、変換 API を順次実装していきます。

本稿では、ARC ポータルデータベースのデータを IIIF マニフェストに変換する API の紹介、データベースから IIIF マニフェストを取り出す方法の紹介を行いました。ARC のデータベースを利用して研究活動を進めておられる方々には、データベース構築活動において、新しく特別なことをしていただく必要はありません。データが公開状態で運用されていれば、という限定は付きませんが、これまでどおり、日常の研究活動の結果をデータベース上に蓄積していただくだけで、ARC ポータルデータベースの相互運用性や、収録される資料の活用可能性が高まっていきます。以上の機能が ARC のポータルデータベースに備わっていることをご理解いただき、ARC のデータベースを活用した研究展開に取り組んでいただければ幸いです。

(以上、文責 金子貴昭)

<ARC Database Basics: Part 2>

◇How to Generate the IIF Manifest from the ARC Portal Database◇

1. The Spread of the IIF
2. The Dilemma of Interoperability and Research Databases
3. An API to Generate IIF Manifest Automatically

1. The Spread of the IIF

A long time has passed since there were calls for standardization and interoperability of metadata. It has been almost 20 years since Dublin Core became known as a metadata convention in the world of humanities research. In recent years, the International Image Interoperability Framework (IIIF) has gained prominence and further become a widely-used international interoperability framework for digital images.

Under the IIIF, a setup file called the IIIF Manifest, which describes general information about the content (metadata and image information), is recognized by IIIF viewers such as Mirador and Curation Viewer to browse the contents. From the user's point of view, the benefits of this approach are enormous as they no longer have to learn how to search individual databases operated by each holding institution and their image viewing methods and can browse materials in a standardized manner.

2. The Dilemma of Interoperability and Research Databases

For over twenty years since its establishment in 1998, the Art Research Center (ARC) has been digitizing cultural resources held by institutions in both Japan and abroad, including the ARC itself, as well as accumulating them as research materials and building their databases. While the databases serve as catalogs of the ARC's holding materials, of course, and provide their information to external users, since its inception, the ARC as a research institution has been developing and accumulating databases to create databases for researchers. In other words, the ARC has been developing databases with metadata structure from the researcher's perspective and with database functions that can be regarded as research tools, without seeking to make them easy to understand for the general public.

By nature, researchers are enthusiasts in their fields. Thus, the more the researcher's perspective is stressed in database development, the more the metadata will incorporate the enthusiast's point of view, which will inevitably veer away from standardization and interoperability, necessary to overcome the boundaries of the field. Of course, this may be less problematic if the research can be completed with a single database. However, it has become essential to publish digitized cultural resources online, conduct research activities while subsuming digital cultural resources in related fields, and curate

them online. In this context, standardization and interoperability need to be ensured. On the other hand, it would place a heavy burden on humanities researchers to learn the IIIF philosophy and functions from scratch and reconstruct metadata structures and databases accordingly while furthering their research. In other words, it would not be practical to ask the researchers to do so.

3. An API to Generate IIIF Manifest Automatically

In view of the above, the ARC's Technical Support Board has developed and started implementing an API that automatically generates the data stored in the ARC databases to the IIIF manifest (hereinafter referred to as 'Conversion API'). In the ARC portal databases, each record has its permalink (permanent link), which is a fixed link to an individual record. In the Ukiyo-e Portal Database, for example, there is a permalink at the bottom of the detailed information display screen:

<https://www.dh-jac.net/db/nishikie/arcUP0001/portal/>

If you append the character string 'manifest.json' to the end of the permalink, it becomes its IIIF manifest as follows:

<https://www.dh-jac.net/db/nishikie/arcUP0001/portal/manifest.json>

In the case of the Early Japanese Books Portal Database, you can find the permalink at the bottom of the detailed bibliographic screen of each volume:

<https://www.dh-jac.net/db1/books/arcBK01-0001/portal/>

If you append 'manifest.json' to the end of this, you will have its IIIF manifest:

<https://www.dh-jac.net/db1/books/arcBK01-0001/portal/manifest.json>

When opening the IIIF manifest in a web browser, you will see that it contains standardized metadata and information, such as the storage location of the images, in the machine-readable JSON format. By specifying this IIIF manifest (manifest URL) in the IIIF viewer described above, you can browse the appropriate materials even if you are not in the ARC web database, which means that the interoperability could be secured.

As for data open to the public that are within the ARC's portal databases, the ARC has been working with JAPAN SEARCH (<https://jpsearch.go.jp/>), a cross-disciplinary integrated portal operated by the National Diet Library. Since the ARC also provides JAPAN SEARCH with the data's IIIF manifest, you can browse the materials included in the ARC's portal databases on the platform of JAPAN SEARCH.

Apart from such a browsing function, we are currently also developing an online research tool called 'Kinukake' which we have announced at various ARC events (scheduled to be open to the public in FY 2021). As to be reported in more detail in a future issue of this journal, 'Kinukake' is designed to be a research space where users can read in digital cultural resources online, annotate and comment on individual resources, group them, and relate them to each other in an intellectually productive way. The

IIIF manifest will also be used to import digital cultural resources to 'Kinukake', which can function as a born-digital research environment where you can collect research materials, and then analyze and discuss them.

At the time of writing this paper, we have implemented conversion APIs in the following databases respectively.

- ARC Ukiyo-e Portal Database
https://www.dh-jac.net/db/nishikie/search_portal.php
- ARC Early Japanese Books Portal Database
https://www.dh-jac.net/db1/books/search_portal.php
- ARC Modern Japanese Books Portal Database
https://www.dh-jac.net/db1/mbooks/search_portal.php

We will also implement conversion APIs in other core portal databases and separate specialized databases in due course.

This paper has introduced the API for converting data from the ARC portal databases into IIIF manifests and demonstrated how to generate IIIF manifests from the data. Those who are using the ARC databases to further their research activities do not need to start anything special for generating IIIF manifests while building your database. If you just keep on accumulating the results of daily research activities in your database, as always, that will enhance the interoperability of the ARC portal databases and the possibility that the materials you included would be used, provided that the data is operated in a public manner. We sincerely hope that you will take notes on these features of the ARC portal databases and advance your research by utilizing them.

(Responsibility for this article: Takaaki Kaneko; and translated by Emily Li, Biligsaikhan Batjargal, and Keiko Suzuki)