

慶應義塾大学メディアセンター

デジタルコレクション - 公開からこれまでの歩み -

いな き 稲木
りょう 竜

(メディアセンター本部主任)

1 はじめに

2017年4月に「慶應義塾大学メディアセンターデジタルコレクション」¹⁾ (以下「デジコレ」とする)の公開を開始してから5年が経つ。デジコレでは、サイトの公開以降、継続的なコレクション資料の追加公開や他システムへのデータ連携を進めてきた。本稿では、デジコレの体系的な紹介と立ち上げから現在に至るまでに実施してきた取り組みについて報告する。

2 デジタルギャラリーからデジタルコレクションへ

メディアセンターでは、2006年より三田メディアセンターが所蔵する貴重書や特殊コレクションをデジタル化した資料を公開するWebサイトとして「慶應義塾図書館デジタルギャラリー」^{2) 3) 4)} (以下「デジギャラリー」とする)を公開していた。デジギャラリーでは福澤諭吉の著作、高橋誠一郎浮世絵コレクション、インキュナブラコレクション、小山内演劇絵葉書のデジタル化資料が公開されており、インターネットを通じて慶應の貴重資料を誰でも閲覧することが可能なサービスとして重要な役割を果たしていた。また、同じく慶應のデジタル化資料を公開する「KOARA-A」というWebサイトで、信濃町メディアセンターが所蔵している解体新書や富士川文庫などの古医書の資料を公開していた。デジギャラリーは公開開始から約10年が経過し、体系的にいくつかの課題を抱えていた。そこで、2016年度より、デジギャラリーに代わる、新たなデジタルアーカイブ公開システムの検討を進め、開発を開始することとなった。

(1) デジギャラリーの課題

長年の運用のなかで、デジギャラリーは次に挙げる体系的な課題を抱えていた。

・機能拡張性が低いシステム

デジギャラリーのWebサイトは、CMS⁵⁾のようなコンテンツ管理システムをベースにしていなかったため、コンテンツの追加や登録済のメタデータの修正を容易に行うことができないという問題があった。また、デジタル画像の公開方式について、標準の規格に則ったシステムではなかったため、外部サービスとのシステム連携を広げていくことが困難であった。

・古い技術を使用していた

デジギャラリーでは、Webブラウザでデジタル資料を閲覧するビューアを使用するためにFlash Playerという動作環境が必要であった。しかし、このFlash Playerは2020年までにソフトウェアのサポートが終了となることが決まっており、以降はWebブラウザを使用して資料を閲覧することができなくなる状況であったため、別のビューアへのシステム変更を進めていく必要があった。

以上の問題点を踏まえて、後にデジコレとして運用することとなる新たなデジタルアーカイブ公開システムのコンセプトを検討していくこととなった。

(2) デジコレのコンセプト

デジコレの開発に当たって掲げたシステムのコンセプトは以下のとおりである。

・公開サイトの統合

前述のとおり、これまではメディアセンターでデジタル化した資料は「デジギャラリー」と「KOARA-A」に分散して公開されている状態であった。これらを統合して1つのサイトで公開するようにした。また、三田メディアセンターが所蔵する奈良絵本・絵巻コレクションや、1997年に慶應義塾大学HUMIプロ

ジェクト (Humanities Media Interface Project)⁶⁾ でデジタル化が行われた、グーテンベルク42行聖書を新たに追加した。公開サイトを一本化することで、利用者は1つのWebサイト上でメディアセンターが公開しているデジタル化資料の全体像を見ることができるようになった。

・国際規格への対応

国際的な画像データの相互利用規格であるIIIF (International Image Interoperability Framework : トリプルアイエフ)⁷⁾ に対応したシステムにすることとした。標準的な規格に則ることで、IIIFの規格に対応した外部のシステムとの連携が特別なシステム改修なしに柔軟に行えることが期待できる。また、システムに組み込むビューアもIIIFの規格に対応しているものを採用することとした。IIIFについての詳細は、後の章で触れたい。

・コンテンツ追加・メンテナンスの簡易化

CMSをベースとしたシステムにすることとした。CMSを使用することで、管理画面の機能を使用して、デジコレに搭載しているメタデータの情報の追加や編集やWebページの追加を容易に行うことができるようになり、サイトの運用や開発において利便性が向上した。

・PDF版の公開

Webブラウザでの画像閲覧だけでなく、図書資料については閲覧のしやすさを考慮し、PDF版でも資料を公開することとした。一部の資料ではPDF上に本文の透明テキストを付加したものを公開しており、PDF上で本文テキストの検索を行うことが可能となっている。

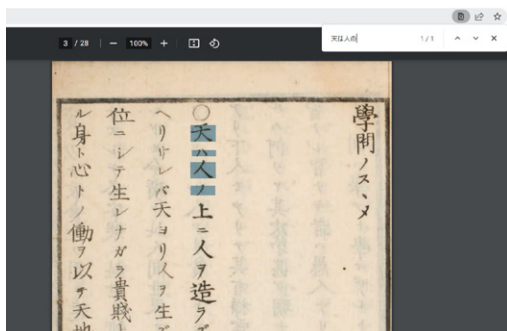


図1 透明テキスト付PDFでのテキスト検索

・検索機能の拡張

デジギャラの検索機能では一部のコレクションの資料のみしか検索できなかったが、デジコレでは搭載されている資料全体を検索できるようにした。搭載されている資料全てのメタデータの検索 (コレクション内の資料検索と全コレクションの統合検索) と、「デジタルで読む福澤諭吉」の資料については本文データの全文検索が可能であり、検索結果から該当のページのデジタル画像を閲覧することができる。

3 デジコレのシステム構成

デジコレのシステムはWebページを公開するためのWebサーバと、大容量の画像データを保存し、配信するための画像サーバの2つで構成されている。また、前述のコンセプトの機能を実現するためにCMSや検索エンジン、IIIFの規格に対応したビューアや画像配信システム等のいくつかのオープンソースソフトウェアを組み合わせて実装している。以下にシステム構成の概略図を示す。

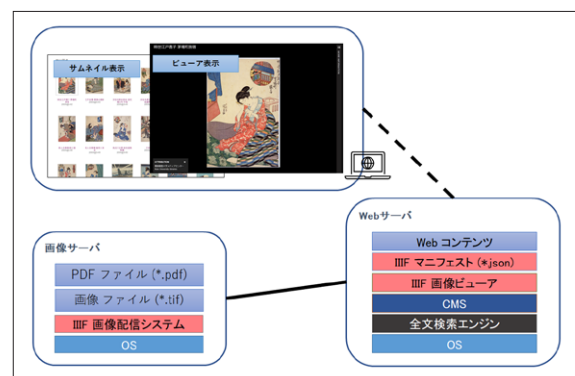


図2 デジコレのシステム構成 概略図

4 IIIFについて

デジコレは、デジタル画像配信の方式として、国際的な標準となっているIIIFに対応している。デジタル画像を相互利用するための枠組みである。IIIFに対応することで、デジコレで公開している画像データは外部のシステムと柔軟に連携したり、人文学研究において、デジタルデータとデジタル技術を活用した、デジタル・ヒューマニティーズの研究に利用したりすることが可能となっている。

(1) デジコレのIIIFへの対応

IIIFに対応するためには、IIIF Presentation API⁸⁾

に準拠したIIIFマニフェストファイル（以下「マニフェスト」とする）という、デジタル資料の構成情報を定義したファイルを用意する必要がある。このマニフェストをIIIFに対応したビューアに読み込ませることで、手元に画像ファイルがなくてもどこからでもビューアで画像を閲覧することが可能となる。デジコレでは、公開している資料すべてについてマニフェストへのリンクを提供しており、誰でも容易に利用することができる。



図3 デジコレで公開しているIIIFマニフェスト

また、デジコレの画像サーバはIIIF Image API⁹⁾に対応した画像配信システムとなっており、IIIF Image APIで定められているURLのリクエスト方法により、画像の大きさを指定して表示させたり、画像の一部を切り出して表示させたり、他にも様々な表示方法を指定して画像を表示させることが可能となっている。

(2) IIIFの仕組みを使用したデジコレの活用例

IIIFの仕組みにより、デジコレ上の資料が多彩に活用されている事例を紹介する。

・富士川文庫デジタル連携プロジェクト試行版¹⁰⁾ (以下「デジタル富士川」とする)

デジタル富士川は、2018年9月より、京都大学図書館機構と共に実施しているプロジェクトである。両館で所蔵する富士川文庫資料をIIIFの仕組みを利用し、1つのWebサイト上に統合して公開することにより、日本医史学の研究・発展への寄与を目的とした取り組みである。2019年2月からは東京大学大学院教育学研究科・教育学部図書室が所蔵する富士川文庫資料もデジタル富士川に連携され、閲覧することが可能となっている。

・顔貌コレクション¹¹⁾

このコレクションは、人文学オープンデータ共同利用センターが公開している美術作品に登場する顔の部分の画像をIIIFの仕組みを使用して切り出し、収集して公開することで、美術史研究に活用できるようにする取り組みである。デジコレの奈良絵本・絵巻コレクションの画像がここで使用されており、切り出した顔画像に情報をタグ付けして公開されている。

5 公開後の取り組み

デジコレは、2017年の公開以降も資料の追加公開や他システムへのデータ連携などを進め、システムの発展に取り組んでいる。これまでに実施してきた主な取り組みをまとめて紹介する。

表1 デジコレのこれまでの取り組み

年度	主な取り組み
2016	・デジタルコレクション サイト構築を開始
2017	・デジタルコレクション サイト公開を開始 [2017.04~] 初期公開コレクション - デジタルで読む福澤諭吉 - インクナブラコレクション - 高橋浮世絵コレクション - クーテンヘルク42行聖書 - 奈良絵本・絵巻コレクション - 小山西内演劇絵葉書 2017 - 解体新書 ほか (解剖学コレクション) - 富士川文庫 (古医書コレクション) ・全コレクションの横断検索機能を追加 [2017.07~] ・公開コレクションの追加 追加公開コレクション - 福澤遺墨コレクション [2017.07~] - 重要文化財 [2018.03~] - 相良家文書 [2018.03~] - 対馬宗家文書 [2018.03~]
2018	・富士川文庫デジタル連携プロジェクト試行版へのデータ連携を開始 [2018.09~]
2019	・公開コレクションの追加 [2019.04~] 追加公開コレクション - ボン浮世絵コレクション - 国書コレクション - 漢籍コレクション ・国立国会図書館サーチ (NDLサーチ) へのデータ連携を開始 [2019.06~] ・ジャパンサーチへのデータ連携を開始 [2019.08~]
2020	・Webサーバ/画像サーバのリプレースを実施 [2020.08~] ・サイトの常時SSL化に対応 [2020.08~] ・公開コレクションの追加 [2020.08~] 追加公開コレクション - 貴重書展示会図録
2021	・Keio Object Hubへのデータ連携を開始 [2021.04~] ・公開コレクションの追加 [2021.11~] 追加公開コレクション - 荒俣宏旧蔵 博物誌コレクション ・PRRLA (Pacific Rim Library) へのデータ連携を開始 [2021.12~]
2022	・公開コレクションの追加 [2022.05~] 追加公開コレクション - 論語疏 巻六 - ピックアップコレクション

※既存コレクションへの資料追加搭載については記載を省略している。

(1) 継続的な資料追加の実施

デジコレに順次追加搭載し公開しているのは、三田メディアセンターのスペシャルコレクション担当が選定した資料のうち利用希望が多い貴重資料と、信濃町メディアセンターでデジタル化が完了している富士川文庫の資料である。公開開始当初は、8コレクション、約2,900件であったが、2022年5月現在では19コレクション、約7,800件まで公開コンテンツが拡大している。また、2020年度からは過去に開催された慶應義塾図書館貴重書展示会の図録の公開を開始した。販売用冊子の在庫がなくなった図録のうち、PDFでの公開が可能なものを順次搭載し、デジコレ上のデジタル化資料と図録の出品リストとのリンク付けをして公開している。

(2) 他システムへのデータ連携の推進

デジコレでは、いくつかの外部システムへデータ連携を行い、外部のサービスからもデジコレの資料情報を検索することが可能となっている。

a 国立国会図書館のシステム

2019年6月より国立国会図書館サーチ¹²⁾（以下「NDLサーチ」とする）へのデータ連携を開始し、NDLサーチ上でデジコレに搭載されている資料を検索できるようになった。また、2019年8月からは、NDLサーチに連携したデジコレのデータは、国立国会図書館が運営するデジタル・アーカイブ・ポータルであるジャパンサーチ¹³⁾にも連携されるようになり、ジャパンサーチ上でもデジコレの資料が検索できるようになった。補足となるが、ジャパンサーチへのデータ連携は2021年4月からは後述のKeio Object Hub¹⁴⁾経由でのデータ連携に変更しており、NDLサーチ経由での連携は停止している。（デジコレからNDLサーチへのデータ連携は現在も継続して実施している。）

デジコレとNDLサーチのデータ連携を開始するにあたり、以下の点について検討し調整を進めた。

・データの利用条件について

コンテンツの利用条件は、デジコレのサイトに掲載しているとおり、画像利用時には申請を必須とし、所蔵情報の表示を条件としている。これは、メディアセンター側がコンテンツの利活用実態を把握した

ということと、申請なしの営利利用を避けたいためである。また、メタデータの利用条件は、CC0¹⁵⁾として二次利用可能な形で公開しており、NDLサーチのAPI提供対象データプロバイダー一覧¹⁶⁾にメタデータの利用条件を掲載している。

・メタデータの連携方法について

デジコレはメタデータを外部へ配信するAPI等の仕組みを有していないため、NDLサーチへのメタデータ連携は、定期的にデータファイルを提供して行っている。具体的には、コンテンツを追加する毎にメタデータをDC-NDL¹⁷⁾のメタデータ記述にマッピングしたデータを作成してExcelファイル形式で提供し、NDLサーチに搭載してもらっている。

b Keio Object Hub

Keio Object Hub（以下「KOH」とする）は慶應義塾ミュージアム・コモンズ（以下「KeMCo」とする）が運営している、慶應が蓄積してきた文化財コレクションを発信するポータルサイトである。KOHへのデジコレのデータ連携は2021年4月のKOHの公開当初から開始しており、KOH上でもデジコレの資料を閲覧することができる。KOHとのデータ連携は前述のNDLサーチへの連携と同じ形式で、定期的なファイル提供にて行っている。

c PRRLA¹⁸⁾ : Pacific Rim Library¹⁹⁾

Pacific Rim Library（以下「PRL」とする）は、メディアセンターが2020年7月に加盟したPRRLA（Pacific Rim Research Libraries Alliance）が実施しているプロジェクトである。PRRLAに加盟している40以上の環太平洋地域の大学が所有しているデジタルリソースを集めて世界に公開し、研究や教育に活用できるようにするという取り組みである。デジコレでは2021年12月よりデータ連携を開始した。これにより海外からもデジコレの利用が広がることが期待できる。PRLへのデータ連携はOAI-PMH（Open Archives Initiative Protocol for Metadata Harvesting）²⁰⁾のメタデータ収集プロトコルに準じたXML形式でのデータ提供が必須となっていたため、所定の形式でデータ提供を行える仕組みを開発し、データ連携を可能にした。

6 おわりに

デジコレは、外部のシステムへのデータ連携やIIIFへの対応により、資料の発見や利活用の幅が広がったといえる。デジギャラで公開していた時と比べ、資料の利用申請数は増加しており、コンスタントに申請が来ている状況である。

今後も継続的なコレクション資料の追加を進め、公開資料を充実させていきたい。また、KOSMOS²¹⁾との連携や今後も新たな機能が追加されていくであろうIIIFの最新バージョンへの対応なども検討し、さらにシステムを発展させていければよいと考えている。

注・参考文献

- 1) 慶應義塾大学メディアセンターデジタルコレクション.
<https://dcollections.lib.keio.ac.jp/>, (参照 2022-06-08).
- 2) 倉持隆. 貴重書業務とデジタル化の展開. MediaNet. 2007, no. 14, p. 26-28.
- 3) 原田奈都子. よみがえる福澤諭吉 — デジタルで読む福澤諭吉 —. MediaNet. 2008, no. 15, p. 36-37.
- 4) 徳永聡子. デジタルギャラリー「インキュナブラコレクション」— 図書館からの国際発信に向けて —. MediaNet. 2008, no. 15, p. 42-43.
- 5) コンテンツ・マネジメント・システムのこと。Web制作に必要な専門的な知識が無くても、Webサイトやコンテンツを構築・管理・更新できるシステム。
- 6) 櫻村雅章. 連載: HUMIプロジェクトの貴重書デジタルアーカイブ (第1回) HUMIプロジェクトについて. 情報の科学と技術. 2006, vol. 56, no. 4, p. 183-187.
- 7) International Image Interoperability Framework.
<https://iiif.io/>, (accessed 2022-06-08).
- 8) Presentation API 2.1.1. International Image Interoperability Framework.
<https://iiif.io/api/presentation/2.1/>, (accessed 2022-06-08).
- 9) Image API 2.1.1. International Image Interoperability Framework.
<https://iiif.io/api/image/2.1/>, (accessed 2022-06-08).
- 10) 富士川文庫デジタル連携プロジェクト試行版.
https://www.kulib.kyoto-u.ac.jp/rdl/digital_fujikawa/index.html, (参照 2022-06-08).
- 11) 顔貌コレクション. 人文学オープンデータ共同利用センター.
<http://codh.rois.ac.jp/face/>, (参照 2022-06-08).
- 12) 国立国会図書館サーチ.
<https://iss.ndl.go.jp/>, (参照 2022-06-08).
- 13) ジャパンサーチ.
<https://jpsearch.go.jp/>, (参照 2022-06-08).
- 14) Keio Object Hub.
<https://objecthub.keio.ac.jp/>, (参照 2022-06-08).
- 15) クリエイティブ・コモンズ・ライセンスのCC0.
<https://creativecommons.org/publicdomain/zero/1.0/deed.ja>
- 16) API提供対象データプロバイダー一覧. 国立国会図書館サーチ.
<https://iss.ndl.go.jp/information/api/dpid/>, (参照 2022-06-08).
- 17) メタデータ記述に用いる語彙の国際標準であるDublin Core (ダブリン・コア) を基に、国立国会図書館独自の拡張を加えた形式。
- 18) PRRLA.
<https://pr-rla.org/>, (accessed 2022-06-08).
- 19) Pacific Rim Library.
<https://pr.library.ucla.edu/index>, (accessed 2022-06-08).
- 20) データの自動収集によってメタデータを交換するためのプロトコルの名称。
<https://www.openarchives.org/pmh/>
- 21) 慶應義塾大学メディアセンターの図書館資料検索システム。
<https://search.lib.keio.ac.jp>