

早稲田大学図書館における貴重資料のデジタル化の取り組みについて ～これまで・現在・そしてこれから～

ティムソン ジョウナス（資料管理課）
山本 さぎり（資料管理課）

1. はじめに

2005 年 4 月、早稲田大学図書館は、所蔵する貴重書や古書資料の書誌データおよび全文画像データを作成する「古典籍総合データベース¹⁾」のプロジェクトを開始した。当プロジェクトは、あらゆる時代や分野にまたがる貴重資料の網羅的なデータベースの構築として結実し、蓄積されたデータは当館の蔵書検索 WINE から検索できるようになった。また、「洋学コレクション」、「大隈重信関係資料」といった、テーマ別ギャラリーを公開し、WEB 展示会形式で各種資料の解説を付することを可能としたほか、国内外のあらゆる利用者が、研究目的であれば手続きを経ず、容易に画像データ（JPEG および PDF 形式）を利用できるようにすることも可能にした。当初 5 ヶ年の予定で開始された当プロジェクトは、5 年目の終了時点で約 90 万件ものアクセスを計上し、国内における総合的な古典籍資料データベースの先駆けとなった。プロジェクトが立ち上がってから 17 年経過した現在も、この古典籍資料のデジタル化事業は継続しており、様々な貴重資料の書誌および画像データが順次公開されている。

2. International Image Interoperability Framework

当館が古典籍総合データベースから取得可能にしていた JPEG や PDF ファイルは、資料の全文を読む、資料のイメージを得る、という面だけでいえば、非常に有用なものであったといえる。しかしながら、例えば複数機関に所蔵されている資料画像を比較検証したいといった研究者からすると、各機関のデジタルアーカイブ（画像リポジトリ）を行き来して各画像を取得し、それらを複数の画像ソフトの画面を立ち上げた上で並べて確認しなければならないなどの手間がかかる問題があり、当館をはじめ、各館でデジタル画像のプラットフォームを構築することが研究の利便性に十分に寄与しているとは必ずしも言えない状況があった。

2010 年代に入り、上記のような問題を解決すべくデジタルアーカイブ間で画像データの相互運用を可能とすることを目指した国際的な枠組み International Image Interoperability Framework (IIIF)²⁾ がスタートした。各機関がこの IIIF 規格に対応した画像、および資料の構造や書誌情報、画像の利用条件や資料の権利情報など

の情報を記述したマニフェスト (Manifest) と呼ばれるファイルを公開することで、研究者をはじめとする利用者は、一つのビューアを通して、複数機関のデジタルアーカイブの画像を利用することができるようになる。

現在、国内でも多くの大学がこの IIIF 規格で貴重資料のデジタル画像を公開している。古典籍総合データベースを立ち上げ、貴重資料のデジタル画像の提供を長らく実践してきた我々は、これまでこの IIIF 規格での画像公開に参画することができておらず、いささか後塵を拝する状況が続いていたが、この度、満を持して IIIF 規格でのデジタル化プロジェクトを立ち上げることとなった。

3. プロジェクトの立ち上げにあたって

古典籍資料の画像を IIIF 規格で公開するにあたっては、画像ファイルの形式を Pyramid Tiled TIFF 形式に変換するか、JPEG 2000 を利用することが必要となる。当館では、各画像を Pyramid Tiled TIFF に変換して作業することを基本方針としている。画像の Pyramid Tiled TIFF 化にあたっては、ImageMagick³⁾ などのフリーソフトを使用する方法も考えられるが、当館では IIIF 事業の立ち上げにあたり、Python 等の技術に明るい者が中心となって、独自の画像変換アプリケーションを作成した。

当館に限らない国内の大学図書館の課題として、技術に長けた人材が常に豊富にいる状態ではなく、今後、専門領域の異なる者が IIIF 事業の担当に配置される可能性も十分に考えられる。今後は各種作業において、フリーソフト等の外部ツールを使用すること等も検討が必要になってくるだろう。

なお、画像は高解像度であるほど容量が大きくなるが、IIIF 対応により画像データ閲覧時の負荷軽減が期待できることから、2022 年度から資料撮影時の解像度を引き上げる変更を行った。撮影画質を決めるにあたっては、既に IIIF 対応をしている他機関の画像を参考に、RAW データで画素数 6000 × 4000 ピクセル・解像度 350 dpi としている。

単に画像を IIIF 規格で公開するだけでは、このプロジェクトに取り組む本来の目的を見失ってしまうことになりかねない。例えば、美術的な価値のある艶麗な作

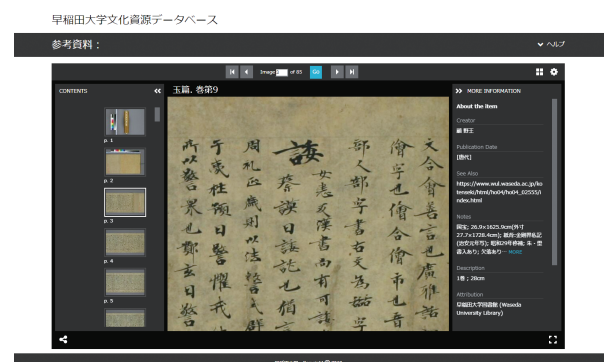
品の高精細画像を IIIF 規格で公開することで、細部まで鑑賞や検分をしやすくするといった貢献の仕方もあるが、IIIF 規格で画像化する資料の対象は、いわゆる絵画に限定されることはなく、手写本や地図、版本など、幅広く存在することに留意する必要がある。その上で、IIIF 規格で公開された資料画像がどのように使われるか、それらを使ってどういった研究が可能かということを考えることが求められるだろう。近年は人文学の研究にデジタル技術を適用、応用する「デジタル人文学」が徐々に盛んになってきている。当館においても、こういった分野の動向について図書館職員が見識を備えておく必要があるということを、プロジェクトの立ち上げにあたり確認した。

4. 実際の運用～画像の公開

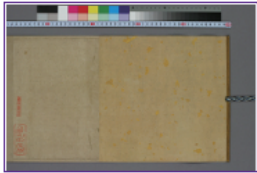
当学には、文化推進部が中心となって管理を行っている「文化資源データベース⁴⁾」があり、学内の博物館や図書館等が所蔵する貴重な文化資源を国内外に向けて発信している。IIIF のための画像サーバもこのデータベースの枠組みで設置されており、各博物館で作成した IIIF 規格の画像は、このデータベース上で公開されている。当館においても、作成した IIIF 規格の画像は文化資源

データベース用の IIIF サーバへアップロードすることを基本としている。当館作成のいくつかの画像は、同データベース内の「バーチャルミュージアム⁵⁾」にて、オンライン展覧会の形で閲覧することができる。

古典籍総合データベースの IIIF 対応も進めている。これまで、撮影した古典籍画像データは、JPEG、PDF ファイルにして掲載しているが、IIIF 対応が可能な画像を作成次第、該当する資料の書誌情報ページに IIIF ビューアの URL およびマニフェスト URL (IIIF ビューアで利用者に提供するために必要な全般的な情報) を追加することになっている。



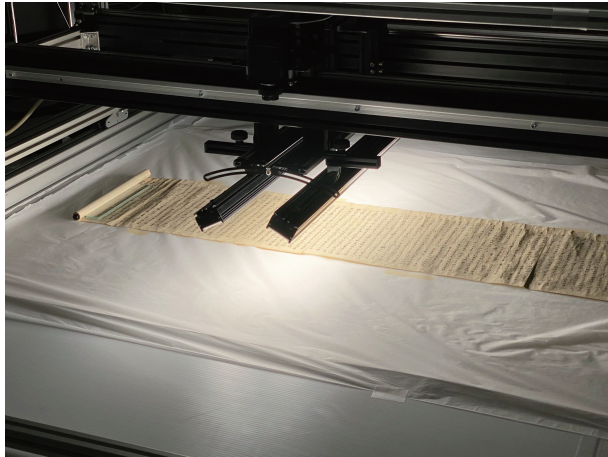
IIIF 対応ビューアを使用して見た『玉篇』の画像

玉篇. 巻第9 / [顧野王] [撰]	
 ↑画像をクリック(Click here)	
ホ4 2555 UV IIIF Image(UV) Manifest	
請求記号 Call No.	ホ04 02555
タイトル Title	玉篇. 巻第9 / [顧野王] [撰] gyokuhen ●このタイトルで早稲田大学蔵書目録を検索(OPAC title search)●
著者/作者 Author	顧野王, 519-581 ko, yaō ●この著者/作者で早稲田大学蔵書目録を検索(OPAC author search)●
出版事項 Imprint	写, [唐代] sha
形態 Description	1巻 ; 28cm
内容等 Notes	国主 26.9×1625.9cm(外寸27.7×1728.4cm) 紙背:金剛界私記(治安元年写) 昭和29年修補 朱・墨書入あり 欠落あり 欠損あり 卷子装 銭屋総四郎,磯淳,川田剛,田中光頭旧蔵 吉部-幸郎

『玉篇』古典籍総合データベースの書誌情報画面。Manifest や IIIF Viewer へのリンクを掲載している

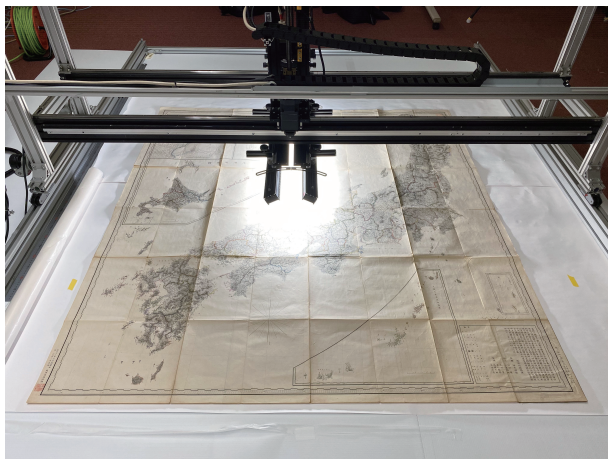
URL: https://www.wul.waseda.ac.jp/kotenseki/html/ho04/ho04_02555/index.html

さらなる高精細画像の公開にも着手している。この春、当館所蔵の国宝、重要文化財および原寸で肉眼では判別しづらい細密な描写がある資料を選定し、株式会社サビア⁶⁾の協力を得て、文化財専用の超高精細スキャナーを用いての撮影を行った。作成された画像データを拡大した様子は、本号の表紙および解説を参照されたい。肉眼では判読できない文字・描写の精密さ・下地の繊維まで見えるほどである。



『礼記子本疏義』撮影の様子

当館が IIIF 規格で公開している古典籍資料の画像の数はまだわずかではあるが、既に学内の研究者から高評価を得ている。先述の通り、細部を確認できるようにすること、他機関が IIIF 規格で公開している画像と一つのビューアで並べて比較できることなどが、世界中の研究発展への寄与につながっているのだと我々自身が実感している。



『大日本絵図』撮影の様子

5. 今後

各機関がマニフェストさえ用意していれば、あらゆる IIIF 規格の画像が任意のビューアで閲覧可能となるのは IIIF の最大のメリットの一つだろう。

例えば海外では、13 世紀フランスのボーヴェ典礼書の写本から切り取られた零葉を所蔵する各機関が協働し、各零葉を IIIF 規格で画像化して一つのビューア上で閲覧できるよう、デジタル“再製本”をする試みが行われている⁷⁾。このコンセプトに従えば、切り取られた卷子物の“復元”や、上下巻が別々に所蔵されている古典籍資料を“揃える”ことなども可能になってくる他、世界中に散逸している愛書家のコレクションを、一箇所に“再構築”することも可能になる。当館は和洋問わず、多数の貴重資料を所蔵している。このような“デジタル複製”のプロジェクトに当館が参画することもあれば、当館が率先してプロジェクトを立ち上げ、他機関に参画を依頼することも十分に考えられる。

また、IIIF 規格で公開した写本や版本の画像に、本文のテキストデータを TEI (Text Encoding Initiative)⁸⁾ 規格でマークアップしていくプロジェクトも国内外で進んでいる。国内においては IIIF 画像に対応したくずし字解読のシステムの開発も進められており⁹⁾、未翻刻資料の解読を支援する仕組みも整ってきている。当館でも、例えば教員や大学院生との協働により当館所蔵の古文書を翻刻し、そのテキストデータを TEI 規格で Omeka¹⁰⁾ などのコンテンツ管理システムに登録して検索可能な形で公開する、といったことを検討していきたい。

IIIF をはじめとするデジタル技術は、従来の人文社会科学の研究手法を補完するものだといえる。デジタル技術が進歩するにつれて、できることもより増えていき、研究手法に幅を持たせることにもなるだろう。

また、図書館が IIIF 規格で資料画像を提供することは、学内外、ひいては国内外の研究者がそれぞれの研究分野の中で、あるいは分野を超えて協働する場やきっかけを作ることにもつながるといえる。この協働をきっかけに学際的な研究が生まれることも期待される。

資料の単純なデジタル撮影を超え、最先端の技術に対応した公開を行うことで、図書館自体の研究支援の幅が大きく広がることが分かった。当館が、当学をはじめ国内外の研究をさらに活性化させるための一つの拠点となるべく、今後もデジタル技術およびデジタル技術を使用した研究の動向を常に意識しつつ、資料の提供、協働研究促進のための情報発信に努めていく所存である。



KuroNet くずし字認識サービス (URL: <https://mp.ex.nii.ac.jp/kuronet/>) にて「つれつれ草上巻」を解析

注

- 1) 古典籍総合データベース
<https://www.wul.waseda.ac.jp/kotenseki/>
- 2) IIIF の WEB サイトで、IIIF 規格の説明やチュートリアルが確認できる。
<https://iiif.io/>
- 3) ImageMagick は多数のファイル形式に対応の画像処理ツール。
<http://www.imagemagick.org/script/index.php>
- 4) 早稲田大学文化資源データベース
<https://archive.waseda.jp/archive/index.html>
- 5) 早稲田大学文化資源データベース内のバーチャルミュージアムの「図書館」のカテゴリに、当館公開の IIIF 規格画像を参照できるコンテンツがある。
<https://archive.waseda.jp/archive/vm-top.html>
- 6) 株式会社サビアは美術品や文化財の高精細スキャニングサービスを提供する会社。
<https://sabia.co.jp/>
- 7) アメリカ中世アカデミーの Executive Director である Lisa Fagin Davis 氏が中心となり進めているプロジェクト。
<https://brokenbooks2.omeka.net/>
<https://fragmentarium.ms/overview/F-yi3b>
実際にデジタル“再製本”した写本は下記のマニフェストを使用して IIIF 画像のビューアで閲覧可能。
<https://fragmentarium.ms/metadata/iiif/F-yi3b/manifest.json>
- 8) Text Encoding Initiative は資料の本文の電子テキストデータを構造化するための国際的な規格を策定している協会。ウェブサイトでは各種資料媒体向けの規格を確認することができる。
<https://tei-c.org/>
- 9) 人文学オープンデータ共同利用センターのウェブサイトで、くずし字の OCR 認識・解読のアプリケーションが紹介されている。
<http://codh.rois.ac.jp/char-shape/app/>
- 10) Omeka はデジタルコレクション向けのフリーのコンテンツ管理システム。
<https://omeka.org/>