

IIIF Curation Platform : 利用者主導の画像共有を支援するオープンな次世代 IIIF 基盤

北本 朝展 (情報・システム研究機構 データサイエンス共同利用基盤施設 人文学オープンデータ共同利用センター/国立情報学研究所)

本間 淳 (フェリックス・スタイル)

Tarek Saier (Albert Ludwigs University of Freiburg)

IIIF Curation Platform (ICP) とは、画像配信基盤として国際的に普及した IIIF (International Image Interoperability Framework) に対して、利用者主導型の画像共有という新たなコンセプトを導入する次世代 IIIF 基盤である。本論文は ICP の中核となる「キュレーション」というコンセプトを紹介し、このコンセプトが ICP を構成するコンポーネントの連携によってどのように実現されているかを説明する。さらに、いくつかの ICP アプリケーションを事例として ICP の活用可能性を議論し、最後に今後に残された課題をまとめる。

IIIF Curation Platform: Next Generation IIIF Open Platform Supporting User-Driven Image Sharing

Asanobu Kitamoto (Center for Open Data in the Humanities, Joint Support-Center for Data Science Research, Research Organization of Information and Systems / National Institute of Informatics)

Jun Homma (FLX Style)

Tarek Saier (Albert Ludwigs University of Freiburg)

IIIF Curation Platform (ICP) is a next generation IIIF platform that introduces a new concept called user-driven image sharing into IIIF (International Image Interoperability Framework), an internationally adopted platform for image delivery. This paper first introduces the concept of “curation” as the core of ICP, and explains how this concept is realized by the cooperation of various components. Moreover, we discuss the applicability of ICP based on case studies on a few ICP applications, and finally summarize challenges that need to be solved in the future.

1. まえがき

本論文は IIIF (International Image Interoperability Framework) の拡張に基づく IIIF Curation Platform (ICP) [1] の概念と設計について説明するとともに、その適用例を紹介する。ICP は人文学オープンデータ共同利用センター (CODH) が構築する、画像研究のための情報基盤である。ICP における最も重要な概念は「キュレーション」である。アートの分野では、キュレーションとはアート作品を対象とした展示企画、すなわちテーマを設定し、テーマにふさわしい作品を集め、会場に配置して、来場者にメッセージを伝える活動を指す。本研究はこの概念を IIIF の世界に広げ、IIIF で公開される画像の中からあるテーマに沿って画像を収集し、整列し、オンライン上で共有する仕組みを構築した。アートの世界におけるキュレーションが、美術作品を制作する提供者主導の世界から、作品の収集と展示から作品に新たな意味や価値を与える利用者主導の世界への転換という役割を担

うのと同様に、IIIF キュレーションは IIIF の世界を提供者主導から利用者主導に拡大する役割を担うものである。

本論文はまず利用者主導型の IIIF の背景を2章で紹介する。続いて3章では ICP を構成するコンポーネントについて、4章ではコンポーネントを連携させるインタフェースについてまとめる。さらに5章では ICP のアプリケーションを3点紹介し、6章では ICP の将来の計画についてまとめる。なお本論文はスペースの都合上、IIIF に関する基礎知識を前提とし、IIIF 用語を説明なく用いる。

2. 利用者主導型次世代 IIIF

2.1 提供者主導から利用者主導へ

IIIF は、画像資源への統一かつ高品質なアクセスの実現を目指した仕様である。ここ数年、世界中の先進的なミュージアムやライブラリが画像公開に利用するようになった結果、画像アクセス方法の標準化が世界規模で進みつつある。

この変化は、HTML (HyperText Markup Language)と HTTP (HyperText Transfer Protocol)による相互運用性の向上が引き起こした、1990年代の情報共有革命と似ている。提供者がウェブサーバを立ち上げ、利用者がウェブブラウザでアクセスすれば、世界中の情報資源が単一のソフトウェアでアクセス可能となった。

しかしこの時代はまだ情報の提供者から利用者への一方向的な流れ (Web 1.0) であった。この状況が大きく変化し始めたのが 2005 年頃であり、ウェブは誰もが発信可能な参加型サービスのプラットフォーム (Web 2.0) あるいは「次世代ウェブ」となった。そして API を介した機械可読データの交換や、複数の API をマッシュアップしたアプリケーションの構築が活発化した結果、ウェブは人間だけでなく機械も活用するプラットフォームに成長を遂げた。この歴史的な変化と同じく、ICP は次世代 IIIF としての特徴を有しており、IIIF の世界を「情報の提供者から利用者への一方向的な流れ」から「誰もが発信可能な参加型サービスのプラットフォーム」へと変える。その中核となる概念がキュレーションである。

キュレーションとは、あるテーマに沿ってコンテンツを収集し、それを適切な順番に並べて、新たなコンテンツとして提示する活動を指す。コンテンツに対する多様な見方を共有することで、コンテンツの新たな解釈を生み出すだけでなく、既存の要素の組み合わせから新たな価値を生み出す派生的な創作物である点からも、これは明らかに利用者主導型の活動である。

我々はキュレーションの考え方を実現するために、IIIF の既存仕様を参考にしつつ独自に拡張した Curation API を提案した[2]。しかし、キュレーションと同様の機能は IIIF の既存仕様である IIIF Presentation API でも実現できるのではないかと、という疑問は当初から存在した。実際のところ IIIF Presentation API をかなり無理な方法で利用すれば、キュレーションに近い機能を実現することも不可能ではない。

しかし上記の次世代 IIIF という観点から考えれば、IIIF Presentation API は提供者主導、Curation API は利用者主導という異なる設計思想を体現したものであると整理することができる。そして設計思想が異なる以上、両者を混同して扱うことは好ましくないとと言える。逆に言えば、キュレーションを IIIF Presentation API で記述するのに無理があるのは、根本的に異なる設計思想を無理に当てはめているからとも考えられる。

2.2 利用者主導の場合の Linked Data 構造

技術的に見れば、IIIF Presentation API と Curation API の違いは、IIIF という linked data 構造に新たなノードを付加する際の権限の問題とまとめることができる。図 1 に示すように、IIIF Presentation API ではすべてのノードへのリンク

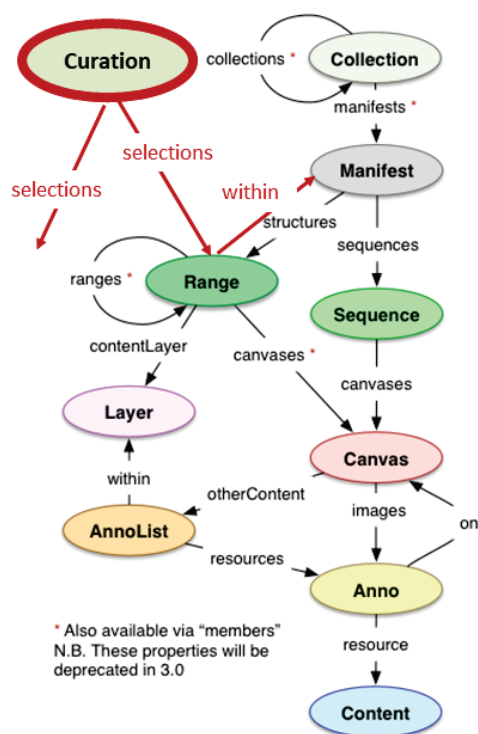


図 1 IIIF Presentation API と Curation API の関係. Curation API は Curation ノードを外付けし、複数の Manifest に対する横断的なリンクを構築できる。

の出発点になるのが Manifest である。IIIF 画像ビューアも通常は Manifest ノードから出発して、そこからたどれるノードを順次読み込んでビューア上に表示する。逆に言えば、Manifest からリンクがたどれないノードは表示できないため、Manifest に新しいノードを付与する権限があるかどうか、情報表示を左右する問題となる。

通常であれば Manifest は画像提供者が用意し、そこにユーザが生成した新しいノードを付加する機能は存在しないため、利用者主導で新しい情報を付け加えることはできないことになる。ならばどうするか。Manifest の内部にノードを付加できないなら、Manifest に向けて外付けで情報を付加していくしかないことがわかる。それがまさに Curation API である。任意の Manifest を対象に外側からのリンクを集約する仕組みが現在の IIIF には存在しない。しかし Curation API を使えば、既存の Manifest を変更することなく、提供者主導型情報と利用者主導型情報をうまく統合できる。これが Curation API の利点である。

現在の IIIF コミュニティは仕様の肥大化を歓迎しないため、今のところ Curation API は ICP 独自の仕様である。しかし我々は、IIIF による正式の仕様化を待たず、独自仕様に対する参照実装を ICP として構築するとともに、機械可読なデータの共有やマッシュアップを通して、より豊かな

IIIFの世界を構築していきたいと考えている。それに向けて、ソフトウェアとアプリケーションの構築を研究の両輪として推進している。

3. ICPのコンポーネント

3.1 ICPの全体像

ICPの最初の成果物である画像ビューア IIIF Curation Viewer (ICViewer) は2016年11月に公開した[2]。しかしICViewerだけではキュレーション基盤としては機能が不十分であったため、他のコンポーネントの構築も順次進めてきた。2017年12月にはJSONkeeperを公開、2018年5月には「顔貌コレクション」(5.1章)を公開し、ICFinderとCanvas Indexerの利用が始まった。そして2018年11月現在、ICPは7個のコンポーネントを有する規模に成長した。7個のコンポーネントは以下のように分類することができる。

1. ICPクライアント (5個)
 - (ア) IIIF画像を閲覧・検索する利用者ツール (3個) ICViewer, ICFinder, ICPlayer
 - (イ) キュレーションを管理・編集する管理者ツール (2個) ICManager, ICEditor
2. ICPサーバ (2個) JSONkeeper, Canvas Indexer

ICPクライアントはプラグイン機構を備えており、必要に応じて機能を拡張できるようにした。またICPサーバは機能を単純化することで、マイクロサービスとして活用できるようにした。またICPサーバやICPクライアントは、HTTP APIで情報をやり取りする。このように細粒度かつ疎結合なシステムとすることで、必要な機能だけを取り出して構成したり、必要に応じて新しい機能を拡張したりすることも可能になる。

さらに複数のツールに対するログインを共通化したシングルサインオンを実現するために、ICPではGoogleが提供するFirebase認証を用いた。Firebase認証の利点は、各種SNSのソーシャルログインやeメールログインに対応している点にある。ICPクライアントのうち、利用者ツールはログインしなくても利用できるが、管理者ツールは基本的にログインが必要である。

なおICPソフトウェアはすべてオープンソースとして公開しており、CODHウェブサイトまたはGitHubからダウンロード可能である。またCODHウェブサイト[1]ではマニュアルやインストールに関するドキュメントも用意しており、本論文よりも詳細な情報はウェブサイトを参照いただきたい[1]。

3.2 ICPクライアント

3.2.1 IIIF Curation Viewer (ICViewer)

画像閲覧機能：IIIF Image APIとIIIF Presentation APIの大部分の機能を実装した、本格的なIIIF画像ビューアである。世界的に見れば、豊富な機能を備えたMiradorとUniversal Viewerのシェアが高いが、ICViewerはこれらのビューアよりも機能

を絞り、シンプルかつ拡張可能なデザインをコンセプトとしている。なおIIIF Image APIへの対応にはLeaflet IIIFを利用し、Leafletコミュニティの豊富なソフトウェア資産も活用している。

キュレーション機能：ICViewerの最大の特徴は、矩形領域を指定して切り取り、キュレーションリストに登録(いわゆる「お気に入り」機能に相当)するというキュレーション機能にある。キュレーションリストはローカルに保存されるが、これをJSONkeeper (3.3.1章)にエクスポートするプラグインを利用すれば、キュレーションを保存、共有することができる。このときキュレーションデータは、Curation APIに従うCuration JSON形式にシリアル化され、これがICPの基本的なデータ形式となる。

メタデータ付与機能：ICViewerのキュレーションメタデータ編集プラグインを利用すると、切り取り画像にメタデータを付与できる。メタデータフォーマットはIIIF Presentation APIを参照し、labelとvalueの配列という極めてシンプルな構造を用いている。ただしlabelは単純なリテラルのため、URIを指定するSemantic Web的な処理はできない。またvalueについては、文字列や数値、配列などいくつかの型を指定できる。

3.2.2 IIIF Curation Finder (ICFinder)

ファセット一覧機能：ICViewerで入力したメタデータの全体像を一覧できるため、キュレーションの傾向を知るために有効である。メタデータのlabelをファセットとし、labelごとにvalueの出現頻度を計測して表示するが、labelとvalueの並び順はAPIを提供するCanvas Indexer (3.3.2章)側で制御する。またvalueはリンクとなっており、同じvalueを共有するキャンバスやキュレーションを検索できる。つまりファセット一覧は検索の入口にもなっている。

検索機能：検索キーワードを入力すると、メタデータの完全一致により検索し、その結果をキャンバスやキュレーションの一覧として表示する。

人間付与・機械付与メタデータ：人間が付与したメタデータだけでなく、4.2章で述べる機械学習サービスとの連携に基づき付与したタグも検索できる。ただし品質の問題があるため、機械付与メタデータは人間付与メタデータとは異なる色を用いて区別して表示する。

3.2.3 IIIF Curation Manager (ICManager)

管理画面：ICViewerまたはICFinderから登録されたCuration JSON形式データを管理するインタフェースを提供する。ICManagerにログインすると、そのユーザが過去にJSONkeeperに登録したキュレーションを一覧できる。この管理画面は、キュレーションの編集、ダウンロード、削除などの機能への入口にもなっているため、キュレーションへのアクセスはICManagerが便利である。

メタデータ編集:ICViewerはキャンパスのメタデータを編集する機能を提供するが、キュレーション自体のメタデータ(タイトル等)は編集できないという制限がある。そこでキュレーションのメタデータは、ICManagerから呼び出すICEditorを用いて編集する。

リスト掲載:ICManagerのリスト掲載/非掲載選択機能は、ICFinderとCanvas Indexerを用いて検索機能を提供する場合のみ重要となる機能である。JSONkeeperに登録するだけでは、キュレーションのIDを共有・公開しない限り、キュレーションに直接アクセスすることはできない。しかし4.1章で述べるようにCanvas IndexerはJSONkeeperの更新通知を受けて、それを検索インデックスに登録してしまうため、検索されることを意図しないキュレーションが公開されてしまう危険がある。そこで検索対象に加えたいキュレーションだけをリスト掲載対象とすることで、公開範囲を制御することができる。

3.2.4 IIIF Curation Editor (ICEditor)

JSON編集インタフェース:JSONkeeperからJSON形式データを読み出し、編集し、再保存するためのインタフェースである。ICViewerには画像を見ながらメタデータを編集できるという利点があるが、画像を見る必要がなければICEditorで直接JSONを編集する方が便利なこともある。またICViewerでは編集できないメタデータはICEditorで編集する。

JSONスキーマチェック:ICEditorは簡易的にスキーマをチェックする機能を有する。あらかじめ設定したJSONスキーマに違反する変更を許可しないことで、破壊的な変更を防ぐことができる。

3.2.5 IIIF Curation Player (ICPlayer)

キュレーションの自動再生:ICViewerでもキュレーションを手動でコマ送りして閲覧することは可能だが、キュレーションの展示や教育への利用では、画像と説明文の同時表示や、自動コマ送りなど、ICViewerとは異なるニーズがある。そこで、ICViewerからキュレーション作成機能やメタデータ編集機能などを削除し、キュレーションの再生に特化したICPlayerを構築した。これは自動コマ送りの速度など、表示方法オプションを変更する機能も備えている。

汎用性:ICPlayerに相当するソフトウェアは、ユースケースに応じて多様なデザインの可能性がありうるため、汎用性には限界がある。例えば5.3章で紹介する高知県立美術館の事例では、要件も実装も大きく異なる状況で、ICPlayerに相当するソフトウェアを独自に構築している。現在のICPで提供するICPlayerは、あくまで基本的な機能を搭載したショーケースとみなすべきであり、

今後も要件に応じて様々なソフトウェアがデザインされることになると考えている。

3.3 ICP サーバ

3.3.1 JSONkeeper

JSON形式データを扱うための汎用的なJSONストアであり、ICPのCuration JSON形式データのストアに限らず、一般的なJSONストアとしても利用可能である。APIはシンプルであり、HTTP POSTでIDを付与した上で、IDを指定したHTTP GET、HTTP PUT、HTTP DELETEをサポートする。このAPI仕様は、Javaで書かれたJSON BlobというオープンソースソフトウェアのAPI [3]を参考にした。ただしJSONkeeperは、JSON Blobと同等の機能をPythonのFlaskで再実装した上で、Firebase認証などの独自機能を付け加えているため、JSON Blob以上の機能を備えている。

JSONkeeperは独自のユーザ認証機構は備えていないが、POST時にトークンを登録しておく、以後のアクセスでは同一のトークンを与えないとPUTとDELETEを不可にするという認可機構を備えている。したがって、このトークンにFirebase認証で得られるトークンを与えれば、疑似的な認証と認可を実現することができる。

3.3.2 Canvas Indexer

キュレーションのメタデータをインデックス化し、検索可能とするためのソフトウェアである。キュレーションとして登録された画像の順序付き集合をバラバラに分解してデータベースに登録することで、メタデータの同一性を用いた画像集合の再編成ができるようになる。そしてメタデータをクエリとして与えると、完全一致するメタデータを含むキャンパスまたはキュレーションの一覧を返すAPIを備えている。

Canvas Indexerの特徴的な機能がファセット一覧機能である。これはメタデータのlabelをファセットとし、labelごとにvalueの出現頻度を計測する機能である。そして人文学分野における要望を反映して、labelやvalueの出現順を細かく制御するための設定も指定できるようになっている。一般にこの種の表示では、頻度順にソートしてしまう場合が多いが、頻度順はデータセットの傾向次第でコロコロ変わるため、研究者はむしろ使いにくいと感じることが多い。さらにこの並び順は、学問分野における重要性などを反映することもあり、特に人文系の学者には強いこだわりがある部分である。そこでCanvas Indexerの設定ファイルではカスタムソート指定を許容することによって、最も使いやすいソート順を細かく制御できるようにした。

3.4 ICP コンポーネントの連携例

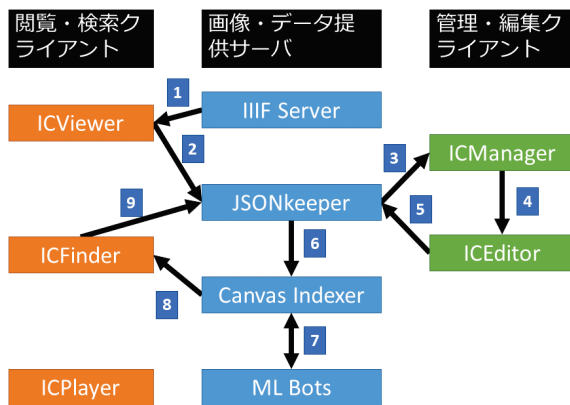


図2 ICP コンポーネントの連携例。

ここまでで紹介した7個のコンポーネントがどのように連携するのか、個別コンポーネントの紹介だけではわかりづらいと考えられるため、5.1章で紹介する「顔コレ」を題材として、各コンポーネントの連携を図2によって説明する。

- 最初に ICViewer が IIF Server からデータを読み込む。IIF Presentation API に従って記述された Manifest を ICViewer にドラッグ&ドロップすると、IIF Image API に準拠したサーバにアクセスして画像を読み込む。
- ICViewer で閲覧しながら、顔などの興味ある矩形領域を切り取って、キュレーションリストに収集していく。次にキュレーションリストの内容を JSONkeeper に保存する。いったん保存したキュレーションを ICViewer で再度表示すれば、メタデータ編集機能も使えるようになる。そこで顔ごとにメタデータを付与しながら、キュレーションを繰り返し上書き更新していく。
- ログインした状態で ICManager にアクセスすると、自分がこれまでに JSONkeeper に保存したキュレーションの一覧を表示できる。すると、キュレーションのタイトルがデフォルトでは「Curated list」となっているため、キュレーションの内容がタイトルでは把握しづらいという問題点に気づく。
- そこで ICManager のリストから、タイトルを変えたいキュレーションに注目し、その行に表示されている編集アイコンを経由して ICEditor に遷移する。
- ICEditor では ICViewer では編集できなかったメタデータも編集できるため、ここでキュレーションの内容に適したタイトルを付与し、JSONkeeper に上書き保存する。
- JSONkeeper にリスト掲載のキュレーションが保存されるたびに、JSONkeeper の Activity Streams が更新される。この更新履歴を定期的に監視している Canvas Indexer は、更新さ

れたキュレーションを取得して、検索インデックスを更新する。

- 画像には人間がメタデータを付与しているが、機械にもメタデータの付与を依頼したい。そこで機械学習ボットと連携し、ボットが画像に自動タグ付けした結果を取得して検索インデックスに加える。これにより、人間には気づかない機械独自の視点によるタグを用いた検索が可能となる。
- ICFinder に利用者がアクセスすると、Canvas Indexer は最新の検索インデックスから生成した結果を返す。ファセット一覧ではメタデータ項目ごとの出現数が一覧できるため、データセットの全体傾向が把握できる。また切り取り画像から画像全体の表示に遷移できるため、どのような文脈で画像が描かれているのかも確認できる。
- ICFinder とのインタラクションを通じて新しい視点を獲得できれば、ICFinder のキュレーションリストに入れていくことで、新たなキュレーションを作成できる。ICViewer とは異なり、こちらのキュレーションは既に切り取られた画像のみがキュレーション対象となる。これを JSONkeeper に保存することで、キュレーションをさらに増やすことも可能となる。

このように、複数のコンポーネントが連携して、キュレーションの作成と編集、共有を実現するのが ICP の仕組みである。ただしすべてを利用する必要がなければ、一部のコンポーネントを省略した構成とすることも可能である。

3.5 ICP Docker 版

ICPのソフトウェアはいずれもオープンソースであるため、自分でダウンロードして設定すれば利用可能となる。このインストール作業は Linux の基礎知識があればそれほど難しくはないものの、一般の人文学者にとって簡単な作業とは言いがたいのも確かである。またサーバ設定には若干ややこしい部分があり、他のソフトウェアとのバージョン依存などの問題も考えると、より簡単にインストールできる仕組みを用意することが望ましい。そこで用いるのが、コンテナ技術として幅広く使われている Docker である。

我々が想定する ICP の使い方は以下の4通りであるが、Docker 版が有用なのは3と4である。

- CODH が提供するデモ環境を利用する場合は、利用の前に用意することは何もないが、プロジェクト専用環境は用意できない。
- ICP クライアントだけを利用する場合は、ソフトウェアを自分のウェブサーバにコピーすれば使えるようになるため、Docker によるインストールは必要ない。ただしこの場合

- も、プロジェクト専用環境は用意できない。
3. ICP サーバを利用する場合は、サーバのセットアップが必要となるが、このサーバ設定は ICP クライアントと比べるとややこしいため、**Docker** を使う価値がある。またこの場合は、プロジェクト専用環境を用意できる。
 4. ICP サーバを利用したいが自力でセットアップすることが困難な場合は、共同研究の一環として CODH が **Docker** 版をホストすることで、プロジェクト専用環境が利用可能となる (5.3 章)。

Docker 版のインストールスクリプトには、ICP の動作に必要な **Docker** イメージをダウンロードし、標準的なサーバ設定を支援する機能があり、**docker-compose** を用いて **Docker** コンテナを 3 つ立ち上げれば、ICP の全機能が活用できるようになっている。

なお **Docker** 版には、IIIF 画像サーバとして **Loris** をインストールする設定も入っているが、IIIF サーバへのニーズは利用者によって違いが大きいことから、デフォルトでは IIIF 画像サーバはインストールしない設定としている。

4. ICP のインターフェース

4.1 JSONkeeper と Canvas Indexer のインターフェース

ICP では複数のソフトウェアが API で連携することで、キュレーションに関連するワークフローを一体的に管理することができる。まず **JSONkeeper** と **Canvas Indexer** との連携について説明する。

ICViewer や **ICFinder** で作成したキュレーションは **JSONkeeper** に保存でき、**JSONkeeper** から読み出せばキュレーションを利用できる。しかし **JSONkeeper** は検索機能を有していないため、キュレーションの検索を実現するには別のエンジンが必要となる。そこで構築したのが **Canvas Indexer** である。

Canvas Indexer はまず、**JSONkeeper** の更新通知をチェックし、新着キュレーションがあればそれをハーベストする。この更新通知では、IIIF **Change Discovery API 0.1** の **conformance level 2** に対応し、**Acvitivity Streams** をベースとする方向の **IIIF Discovery API** の動向を先取りする。そしてハーベストしたキュレーションに含まれるメタデータをインデックス化し、ファセット検索にも対応できるようにする。このようにキュレーションの検索は、**JSONkeeper** と **Canvas Indexer** を **IIIF Discovery API** で連携させるという方法で実装している。

4.2 Canvas Indexer と機械学習サービスとのインターフェース

Canvas Indexer は **ICViewer** や **ICEditor** で人間が付与したメタデータを対象に検索する。しかし最近の機械学習技術の進展を考えると、機械学習による自動タグ付けなど、IIIF に機械学習を活用したというニーズは今後増えてくることが予想される。それを試験的に実現したのが、**Canvas Indexer** と機械学習サービスとを連携させるための API である。

現在のバージョンは、画像の自動タグ付けを機械学習サービスとして想定している。具体的には、ある画像 URL を与えると、それに対して自動的にタグ付けし、タグのリストを返してくれるポットの存在を想定する。5.1 章で述べる「顔コレ」は実際にこの仕組みを用いて、顔貌画像に自動タグ付けを行い、その結果でも画像を検索できるようにした。

この仕組みは原理的には、画像タグ付けタスクだけでなく任意の機械学習サービスと連携させることが可能である。機械学習と IIIF の世界を連携させることで、IIIF の世界に対する多様な処理の可能性を開くハブとなるのが **Canvas Indexer** の役割である。

5. ICP とアプリケーション

ICP は画像を研究素材とする様々な研究に利用できる。**CODH** では ICP のデモを提供しているため、ICP の動作は簡単に試すことができる。しかし研究への利用を考えると、プロジェクト専用の ICP 基盤で他のデータとの混在を防ぎながら、定められたスキーマに従ってメタデータを構築していきたいと考えるのが自然であろう。そこで必要となるのがプロジェクト専用 ICP の立ち上げであり、それは 3.5 節で述べた **ICP Docker** 版を使えば比較的簡単に行える。以下ではプロジェクト専用 ICP を活用した事例を示すことで、ICP がどのような研究に使えるのかを示したい。

5.1 顔貌コレクション (顔コレ)

「顔貌コレクション (顔コレ)」は、絵本や絵巻物から顔だけを切り出してメタデータを付与することで、タグによるファセット検索を実現したものである[4]。

顔コレの初期バージョンは **ICViewer** のみを利用したものだった。顔の切り出しと収集というキュレーション、そして画像の拡大機能による細部比較だけなら **ICViewer** でも可能であり、これによって美術史研究における **ICViewer** の有効性を確かめることができた[5]。さらに **ICViewer** から保存した **Curation JSON** 形式データを **ICViewer** で見られる形で公開することで、研究の再現可能性にも寄与できることを示した[6]。このように、

ICViewer だけを用いるパターンが、最小構成の ICP と言えるだろう。

しかし ICViewer だけではキュレーションの検索ができない点、また全体の傾向が見られないという点に課題があった。そこで我々は ICFinder を構築してこの問題を解決した。新しい「顔コレ」では、画像に付与したメタデータを用いた資料横断的な検索もできる。例えば「正面顔」を全資料から収集して網羅的に比較するなど、従来よりもはるかに大規模なデータに基づく様式研究の可能性が見えてきた。

このように「顔コレ」は細部比較と全体傾向の両方を往來しながら研究を進めていくことが可能な研究基盤となった。これは精読 (close reading) と遠読 (distant reading) を併用した美術史研究を可能とするものとも言える。さらに機械学習を用いた画像自動タグ付けは、人間では思いつかない意外な出会い (セレンディピティ) につながることも見出した。

なお顔コレは 3.4 章で述べたように、ICP のすべてのコンポーネントを活用したアプリケーションである。つまり現状の ICP が実現できる機能をすべて盛り込んだアプリケーションである。

5.2 武鑑全集

「武鑑全集」は、大名家の名鑑である『武鑑』をデジタル的に再編集することで、江戸時代の人や組織、文化などに関する情報基盤を構築する研究[7]である。

武鑑全集では様々な情報を編集しているが、その中に家紋や纏などのグラフィックな要素を切り出し一覧する「紋・道具」というデータベースがあるが、このデータベースに利用するデータを以下の手順で作成した。

1. CODH が提供する「日本古典籍データセット」の中から『武鑑』の IIF 画像を ICViewer に表示し、「紋・道具」の画像を切り取った。
2. ICViewer のメタデータ編集機能で、画像の種類を記述し、その結果を JSONkeeper に保存した。
3. 『武鑑』に登場するすべての大名家に対してこの作業を繰り返した。
4. JSONkeeper の API を用いて、Curation JSON 形式データをダウンロードした。
5. 独自のスクリプトを用いて、画像の種類ごとに画像コレクションを再編成した。
6. 「武鑑全集」のウェブサイト形式に合わせて画像とメタデータを編集した。

この例では、ICP の利用はデータ作成が中心であり、利用者には直接見えない部分での利用が主になっている点が「顔コレ」とは異なる。しかし

ICP のキュレーションデータは機械可読形式であるため、必ずしも ICP でデータ利用を完結させる必要はなく、ツール間の連携やマッシュアップによって利用可能性が広がるという点が ICP の利点である。

5.3 ICP ホスティング

5.1 章および 5.2 章の例は、いずれも CODH が ICP を立ち上げ、自らのプロジェクトを推進した場合である。一方、独自の情報基盤を有さない小規模の美術館や図書館では、ICP を自力で立ち上げることが困難な場合もある。そうした場合への対応として、CODH がクラウド側で ICP をホスティングし、運用を気にせず使える環境を提供する方向での実験も進めている。

その一つが高知県立美術館の事例である。高知県立美術館では 2018 年 10 月 28 日から「芳年」展を開催し、そこに町絵師「絵金」が描いた作品「土佐年中風俗絵巻」の展示システムを設置した。このシステムは、フロントエンドに NPO 法人連想出版が構築した IIF 対応電子絵巻ビューア、バックエンドに ICP を用いており、両者を Curation API で接続する構成とした。

この IIF 対応電子絵巻ビューアは、大きな絵巻を自在に拡大縮小移動し、重要箇所の説明文が読める機能を備えているが、このビューアで利用するデータを以下の手順で作成した。

1. CODH がホストする IIF 画像サーバ (IIFImage) を利用して画像を配信する。
2. ICViewer を用いて、重要な部分の画像を切り取り、キュレーションを作成する。
3. ICManger と ICEditor を用いて、メタデータや説明文などを編集し、JSONkeeper に保存する。
4. IIF 対応電子絵巻ビューアが JSONkeeper から Curation JSON 形式データを読み出して表示する。

ここで ICViewer, ICManger, ICEditor は一つの Docker コンテナ、JSONkeeper が一つの Docker コンテナで動作しており、ICP の一部の機能を利用することでバックエンドを構築した。また Curation JSON 形式データを JSONkeeper 経由で提供することにより、ビューアで表示するデータを遠隔地からでも変更できる利点も生まれた。

このように、ICP という汎用的なバックエンドと、展示ごとにカスタマイズしたデザインによるフロントエンドを組み合わせ、それらを IIF (Curation API) という汎用的なプロトコルで接続することで、一つの展示システムを短期間で構築できたことは、この方法が展示システムとして有効な選択肢になりうることを示す成果である。

6. ICPの将来

ICPは利用者主導の画像共有を促進する次世代IIIF基盤である。従来型IIIFが提供者主導としてミュージアムやライブラリから画像を配信することにフォーカスしたのに対し、次世代型IIIFは利用者がそれらの画像に新しい価値を付け加え、他者と共有することにフォーカスする。

これは人文情報学研究（例えば美術史研究）にも大きなインパクトを与える変化である。なぜなら人文学のデジタル化とは、提供者側のデジタル化、すなわち研究素材となるデータがデジタル化されるだけにとどまらず、利用者側のデジタル化、すなわち研究過程となるワークフローがデジタル化されることも欠かせないからである。資料の加工、選別、注釈、共有といったワークフローがICPの上で、しかも人文学者が使える形になることが今後の重要な課題である。

第一の課題は、各研究分野に特化した比較・計量機能などをプラグインとして実装することで、画像ベース人文学研究で幅広く活用できるようにすることである。第二の課題は、プロジェクト専用ICPを簡単に立ち上げる仕組みである。これは3.4章で述べたICP Docker版や5.3章で述べたICPホスティングである程度は実現できた。

そして第三の課題が、研究段階に応じたオープン・クローズド戦略に対応した認証・認可基盤の構築である。IIIFで認証・認可に関する仕様にはIIIF Authentication APIがあるが、その利用例はまだ少ないのが現状である。しかし人文学分野などでの研究利用を考えると、研究者グループ内でのクローズドな環境で研究を進め、研究が進展した段階でオープン化するための機能には強い要望がある。ICPは一部のツールがFirebase認証に対応しているものの、IIIF Image APIなどの画像配信からキュレーションの共有まで、プロジェクトの全体を一体的にシングルサインオン化することはまだできない。こうした認証・認可基盤の構築は次の大きな課題である。

最後に、ICPの研究を行う動機に触れておきたい。IIIFの世界における問題点の一つに、IIIFに関する固定観念の強さがあると考えている。MiradorやUniversal Viewerなどの普及したツールを使うだけで、高解像度画像を人間が便利に利用できること、それがIIIFの価値であるという説明は決して間違いではない。しかしそれだけで終わってしまうと、IIIFのキーワードであるinteroperability（相互運用性）の価値を存分に活かすことは難しい。相互運用性の価値は、サーバ側とクライアント側で任意のソフトウェアを選んでも、コミュニケーションがきちんと成立することにある。しかも機械可読データのやり取りであ

れば、コミュニケーションの両端は人間である必要さえない。人工知能ボットと対話できるモバイルアプリなどが作れるのも、ウェブの世界に優れた相互運用性が確立しているためである。

それに対して現在のIIIFの世界にはこうした多様性が欠けており、古典的なワークフローを部分的にデジタル化した使い方が多いように感じる。しかし4.2章で述べたようなCanvas Indexerと機械学習サービスとの連携など、今後はワークフローにより多くの機械が導入されることが予想されるし、そこでIIIFの相互運用性はさらなる力を発揮するはずである。

IIIFにおいて固定観念が強いように感じるのは、あるいはIIIFの現在の主力ユーザの考え方が古典的なことが多いからかもしれない。しかしIIIFを他の分野にも広げていくには、より多様性が高いIIIFの世界を実現する必要がある。ICPはそうした方向に向けた一つの試みである。

謝辞

IIIF Curation Platformの美術史研究への活用に向けた共同研究を推進する東京大学の高岸輝氏、CODHの鈴木親彦氏には、美術史研究の推進に寄与する要望の具体化や、キュレーションデータの作成について、多大な協力を受けた。ICPアプリケーションでは、国文学研究資料館や慶應義塾大学メディアセンター、京都大学附属図書館などが公開するデータを利用させていただいた。

参考文献

- [1] IIIF Curation Platform, <http://codh.rois.ac.jp/icp/>, (参照 2018-11-01) .
- [2] 北本 朝展, 山本 和明, "人文学データのオープン化を開拓する超学際的データプラットフォームの構築", 人文科学とコンピュータシンポジウム じんもんこん 2016, pp. 117-124, 2016.
- [3] "JSON Blob API". <https://jsonblob.com/api>, (参照 2018-11-01) .
- [4] 鈴木 親彦, 高岸 輝, 北本 朝展, "顔貌コレクション (顔コレ) 精読と遠読を併用した美術史の様式研究に向けて", 人文科学とコンピュータシンポジウム じんもんこん 2018, 発表予定.
- [5] 鈴木 親彦, 高岸 輝, 北本 朝展, "IIIF Curation Viewer が美術史にもたらす「細部」と「再現性」 絵入本・絵巻の作品比較を事例に", 人文科学とコンピュータシンポジウム じんもんこん 2017, pp. 157-164, 2017.
- [6] "日本古典籍キュレーション". <http://codh.rois.ac.jp/pmjt/curation/>, (参照 2018-11-01) .
- [7] 北本 朝展, 堀井 洋, 堀井 美里, 鈴木 親彦, 山本 和明, "時系列史料の人機分担構造化: 古典籍『武鑑』を参照する江戸情報基盤の構築に向けて", 人文科学とコンピュータシンポジウム じんもんこん 2017, pp. 273-280, 2017.