

6 IIFによるデジタルアーカイブの提供と利活用（ 新技術を活用した新たな図書館サービスの可能性）

著者	永崎 研宣
内容記述	研修：令和元年度大学図書館職員長期研修 主催：筑波大学 期間：令和元年7月1日～7月12日 会場：筑波大学春日エリア情報メディアユニオン2 階メディアホール等
発行年	2019-07
URL	http://hdl.handle.net/2241/00157196

講義題目「IIIF によるデジタルアーカイブの提供と利活用(新技術を活用した新たな図書館サービスの可能性)」

一般財団法人人文情報学研究所 永崎研宣

はじめに

近年、IIIF (International Image Interoperability Framework) ¹という規格が世界中のデジタル画像公開サイト、日本で言うところの「デジタルアーカイブ」の間で急速な広がりを見せつつある。フランス国立図書館やオックスフォード大学図書館、スタンフォード大学、ハーバード大学、ウェルカム図書館など、すでに本格的に採用しているところは枚挙に暇がない。また、いわゆる統合検索サイトである Europeana や DPLA 等でも採用されており、たとえば Europeana では、IIIF を介することで、フランス国立図書館で公開している高精細デジタル画像を自らのサイト上でも直接表示できるようになっている。日本でも、2016年5月にニューヨークで開催された IIIF カンファレンスにて東京大学大学院人文社会系研究科次世代人文学開発センター人文情報学拠点により公表された仏典図像のデジタルアーカイブ、SAT 大正蔵図像 DB²での採用を皮切りに、各地で本格的な採用が広がりつつある。IIIF がこのように急速に広く普及しつつあるのは、Web 上でのこれまでのデジタル画像公開・共有の問題点を抜本的に解決しているからである。

IIIF: 問題とされたのは何だったのか？

IIIF は、Web 上でのこれまでのデジタル画像公開・共有の問題点を抜本的に解決している規格である。では、その問題点とはどういうものだったのだろうか。それを検討していくにあたり、Web 上で発表スライドを共有するためのサイト SlideShare に公開されている、あるプレゼンテーション資料を採り上げてみよう。これは、IIIF の初期からの中心メンバーであるスタンフォード大学図書館チーフ・テクノロジー・ストラテジストの Tom Cramer 氏が2012年にデジタル図書館連盟(DLF)のフォーラムにおいて IIIF の構想を発表した際のスライド(以下、Cramer[2012])である³。これを手がかりとして、IIIF の当初の問題意識の一端を垣間見てみよう。

Cramer[2012]では、デジタル画像の配信を、重すぎ、遅すぎ、高価すぎ、不恰好すぎるとした上で、リポジトリ担当者、ソフトウェア開発者、利用者、資金提供者のいずれもが苦しんでいると指摘する。そして、続くスライドにおいて、幾人かのステイクホルダーの例を挙げる。二つの異なるリポジトリにある写本の筆跡を比較したい古書体学者。様々な資料から教材用の画像をまとめたい美術と建築の教員。歴史的地図の高精細画像に注釈をつけたい人文学研究者しかし彼女

¹ <http://iiif.io/>

² <http://dzkings.l.u-tokyo.ac.jp/SATi/images.php>

³ <http://www.slideshare.net/Tom-Cramer/iiif-international-image-interoperability-framework-dlf2012>

の好む注釈ツールは他のサイトでしか使えない)。高精細ズーム機能付きの新聞画像閲覧ソフトを、開発やカスタマイズをせずに自分のサイトに設置したいリポジトリ管理者。新たな学術資源のデジタル化を支援したいが、改めて一から作ったスタンドアロンのデジタルコレクション Web サイトへの支出はしたくない資金提供者。そのようにして現状のニーズを例示した上で、次に、各地のデジタルリポジトリで公開されている中世写本を例にとり、既存のリポジトリに共通する問題点を挙げる。それは、個別のサイロに閉じ込められて相互運用性がなく、アプリケーションソフトはリポジトリごとに別々に作成されるため開発と維持に費用がかかり、ユーザはほとんど統一されていない多くのインターフェイスになんとかして対応しなければならない、という三点である。そして、この現状を、サイロと重複の世界、すなわち、各地で同じような物が別々に作られてそれぞれが互いに閉じた状態になっている世界である、としている。

欧米の文化機関やデジタル文化資料に関わる人々の間ではそのような問題意識が広く共有されていたようだが、日本では、こういった問題は、まさに「デジタルアーカイブ」という言葉で取り込まれてきたことであり、ここでは、日本のデジタルアーカイブの問題点として捉え直してみたい。それを考えるにあたっては、利用者と提供者という二つの視点でそれぞれに見てみるのがわかりやすいと思われるので、以下に、簡単に見ていこう。

利用者側でのデジタルアーカイブの問題点としては、まず、一つ新たにデジタルアーカイブが構築されるたびに、それぞれが独自性を追求したインターフェイスとなっていることが多いため、その使い方をほぼ毎回新たに覚えなければならないという点がある。同じようできて微妙に違っていることも少なくないので、ストレスなく操作できるようになるのはなかなか大変である。そして、何かを探そうと思ったら、一つずつアクセスし、検索してみなくてはならない。デジタルアーカイブ群の横断検索ができるようになればある程度問題解決できることになるが、結局、横断検索で発見した後、それぞれのデジタルアーカイブに見に行かなければならないので、結局、それぞれの使い勝手を覚えていかなければならない。そして、使い勝手がよくないので改善してもらいたいと思ったとしても、一つの一つのデジタルアーカイブサイトにそれぞれ要求を出さねばならず、しかも後述する理由により、対応してもらえないことは極めてまれである。さらに、見えている画像を自分のサイトで紹介したり、印刷物に取り込んだりしたいと思っても、多くの場合、利用条件が明記されておらず、どうしたらいいのかよくわからない、ということになってしまう。デジタルデータは1と0に還元される単純明快な世界であり、それがこれからの知識基盤となって社会を効率的効果的に支えていくはずなのに、実際には、利用者側からみると、複雑怪奇であり、かつ、非効率的であるといしか言いようがない。このように個々に分断されていることにはそれなりの意味があるのが、しかし、それは利用者の側に立つと十分な説得力があるとは言えないように思ってしまうのである。

一方、提供者側で抱えていた問題点としては、デジタルアーカイブシステムを企業に発注して導入する場合には、いちいちかなり詳細なことまで決めた上で仕様書に記載しなければならないことが挙げられる。仕様書を適切に記載するためには関連する情報の収集をはじめとしてかなりの手間暇がかかってしまう上に、納品検収では仕様書に詳細に書いた内容がきちんと反映されているかどうか確認しなければならず、やはりかなり大変な仕事になってしまう。

あるいは、そのようにしてなんとかして公開したデジタルアーカイブのコンテンツ画像が、他の機関のコンテンツ画像と同じコレクションなので一つにまとめて見えるようにしたい、ということになった場合、あるいは、同じ本の泣き別れがあちらにもあるのでそれを便利に見られるようにしたい、と思った場合、ハードルはさらに高くなる。同じ IT 企業の製品であってもそのようにして相互運用できるとは限らず、まして、異なる企業のデジタルアーカイブシステムを利用していたら、相互運用は諦めざるを得ないということが多いただろう。コンテンツとしての必要性や利用者の利便性とは異なる観点から、そのような制限が生じてしまっているのである。

一方、デジタルアーカイブを公開していると、利用者から改善要求が来ることがある。コンテンツの内容に関することであれば、ある程度対応できることが多いただろうが、画像ビューワの使い勝手などのシステムに関することであれば、多くの公開機関では既製品を使ったり外注したりしているため、改修しなければならないということになり、少なくない費用がかかってしまう。そして、その費用を捻出するために稟議書を回したり会議を何度も開いたりしなければならないこともある。また、ビューワを完全に入れ替えざるを得ないような要求になってしまうと、そもそもまったく対応できないということもある。そのようなことで、利用者の要望に即応するのがなかなか難しいことになってしまっている。

また、公開したところで、そのコンテンツを必要としている利用者のところにうまく届くかどうかという問題が出てくる。近年よく話題になる「横断検索システム」が提供され、そこにコンテンツを載せることができれば、この点についてはある程度解消されることだろう。しかし、従来型の横断検索システムでは、メタデータだけしか検索できなかったり、あるいは、頑張ってもサムネイル画像くらいしか表示できなかったりすることが多く、メタデータが十分に整備されていない場合もあることから、横断検索システムは重要だがそれだけで問題解決するわけではないのではないか、ということも危惧の一つだろう。

このようにしてみると、1 と 0 で構成されているので容易に自由に操作できるはずのデジタルコンテンツが、色々な制約によって、活用しやすい形での公開がなかなかできていないようであることがわかる。ここでは日本のデジタルアーカイブの状況という断りを付けたが、日本に限らず、デジタル文化コンテンツを公開する世界中の多くの機関で多かれ少なかれこのような問題に直面しており、それを多少なりとも解決しなければならないという機運が生まれてきていることが IIF の登場につながってきているのである。

さて、デジタル画像の公開と共有に関してここで指摘されている問題に関しては、我々の多くはおそらく、若干の疑問は持ちつつも、それを所与のものとして受け入れてきたのではないだろうか。そして、状況に対処すべく、個々のサイトがそれぞれに独自のシステムを構築する予算を確保できるように様々な調整を行ったり、個々のサイトの利用方法に精通することを目指して日々研鑽を積んだり、相当な労力をかけて取り組んできた人は少なくないだろう。しかし一方で、たとえば Unicode が導入され普及したことによって世界中のコンピュータで文字コードの相違による相互運用の問題がほとんど解消してしまったことに端的に見られるように、デジタル画像の公開と共有に際しての上記のような問題についても、共通の規格が策定され、それを皆が採用するようになれば、問題の多くは解決してしまうのではないか。そう考えていた人も決して少な

くはないだろう。それでは、これらの問題の解決は、どのようにして行われようとしたのか、そして、実際どのようになっているのか、以下にみてみよう。

IIIF: 課題解決への取り組み

IIIFにおけるこの問題の解決への道筋を辿ってみるなら、やはり、まずはTom Cramerによる上述の発表資料に戻ってみるべきだろう。ここではまず、デジタル画像、おそらくここでの文脈では中世写本のデジタル画像の公開と共有に関する現状のステイクホルダーを分析し、研究者・リポジトリ(管理者)・ツール開発者の間でそれぞれに関心の方向性が異なっていることを提示する。そして、研究者は探索・活用・分析・注釈の最善の組み合わせを求めており、リポジトリは資料の活用・サービスとコンテンツの改善を、ツール開発者は有益なツールやアプリケーションを構築するためにユーザと資料を求めている、とまとめている。

このような現状を踏まえた上で、抜本的な解決策として、画像配信・メタデータ・検索の3つのAPIを定義することが提唱されている。さらに、**Open Annotation⁴**と統合的に利用するために共有キャンバス(Shared Canvas)という仕組みをベースとする、としている。現在利用されているIIIFの基本的な構想は、この時点ですでにほぼ出来上がっていたと言ってもいいだろう。なお、この時点で名を連ねていたのは、フランス国立図書館、英国図書館、コーネル大学、ロスアラモス国立研究所図書館、ノルウェー国立図書館、オックスフォード大学、スタンフォード大学であり、複数の主要な機関が初期の段階から参加していたことは注目に値するだろう。

では、2019年の現状に戻ってみよう。2011年時点で解決すべきとされた問題点の多くがIIIFによって解決できることが示され、体験できるものとして提供されいる。それらがどのように解決されているのか、以下に見てみよう。

IIIFの現状

IIIFが提示した解決策、すなわち、Web上で画像をやりとりするためのルールを共通化してしまうことによって、そのルールに従いさえすれば、誰でも対応するソフトウェアを開発でき、誰でも公開されている画像をアノテーションも含めて自分のサイトで好きなインターフェイス上で表示でき、誰でもそこにアノテーションを付けたりすることができるという状況がもたらされたのである。

そのようになると、利用者は、自分が好きなビューワを1つ選んで世界中のすべてのIIIF対応画像を自由に拡大縮小したり切り出したりしながら見ることができ、それらの画像に対するアノテーションもそこで閲覧・操作することができる。ビューワが気に入らなければ別のビューワを選択しても同じ画像群を閲覧することができる。用途に応じて複数のビューワを使い分けることもできる。どのビューワにも満足できなければ、オープンソースで開発されているいずれかを選んで自分で改良するなり誰かに改良を依頼することもできる。

一方、提供者側では、まず、システムを発注する際に、これまでは、画像のフォーマットや配

⁴ Web上のリソースにアノテーションを付与するための枠組みとして当時提案され徐々に普及していた規格。<http://www.openannotation.org/> 現在はW3CにおいてWeb Annotationとして取り組みが進められている。

信方法からビューワの性能や機能に至るまで、いちいち仔細に仕様書を作成しなければならなかったことが、「IIIF の最新版に準拠すること」と記載するだけでよいことになり、調査や書類作成等にかかる少くないがしかし見えにくいコストの削減とともに提供サービスの品質が相当程度安定化できることになる。また、仕様の策定プロセスがオープンに行われており、仕様策定そのものへの参画の道も拓かれていることから、どうしても必要であれば、仕様を改訂することもできる。さらに、上に述べたように、利用者側でビューワの使い勝手などの問題を解決できるようになることから、提供者側がそこにかかるコストが大幅に削減されることになる。

IIIF は、現在、4つの API で構成されている。画像自体を配信するための Image API、画像やアノテーションを統合して提供するための Presentation API、検索に関するやりとりの仕方を定めた Search API、認証の仕方に関する Authentication API である。特に前二者が現在は広く採用されるようになってきており、個別のサイロに閉じ込められて相互運用性がない、という問題点は、技術的にはこの両者によってほぼ解決可能となっている。そこで、これらについて、以下に少し見てみよう。

Image API について

Image API では、Web 上に公開された画像は、好きなように取り出して各地のサイトに表示することを可能にする。下記のように、サイズの変更や切り出し方の変更等に関する URL の記法が定められており、この API 単体でも画像切り出しに関わる様々な用途に活用でき、他のサイトから画像の一部分のみを直接利用するといったことも可能となっている。

scheme は http か https かを記述するのが一般的である。server は画像を置いているサーバのホスト名 (FQDN)、prefix は画像の置かれるディレクトリ等を記述し、identifier で個々の画像の ID を指定することになる。そして、region 以降が、画像サーバに対して URL で指定する画像変換の内容ということになる。region は画像内での切り出したい画像の位置情報であり、切り出しをせずにそのままが良い場合は full と指定する。size は、切り出された画像 (あるいは切り出さなければ画像全体) のサイズを指定する。たとえばサムネイル画像がほしい場合にはこれを小さな値に設定することになる。rotation は画像の回転度であり回転させない場合は 0 を指定する。quality は通常は default だが、グレー (grey) や白黒二値 (bitonal) の指定ができることになっている。format は画像形式だが、これは jpg とすることが多い。画像形式に関しては、画像サーバとビューワの両方が対応している画像形式である必要があり、筆者がこれまで見てきた範囲では、ほとんどは jpg であった。

```
{scheme}://{server}/{prefix}/{identifier}/{region}/{size}
/{rotation}/{quality}.{format}
```

このように、Image API だけでも、すでに、画像がサイロに閉じ込められていること、開発と維持のコスト、不統一なインターフェイス、といった問題点はある程度解消されている。さらに、それを各サイトで実装するためのソフトウェアもフリーのものが様々に公開されて選択肢も広がっている。Image API に関しては、ICC プロファイルを前提としていないことなど、今後の改善

の余地はあるにせよ、現時点の Web 環境を前提とした場合、十分に枯れており汎用性は高いと言っただろう。

Presentation API について

Image API によって、画像自体の扱いに関しては、かなり色々なことができるようになっていくことがわかる。しかしそれだけでは、画像に付された注釈をどうするか、複数の画像を集めた本や資料としてのまとまりをいかにして効率的に運用するか、といった点は解消できていない。ここに、Presentation API の役割が出てくることになる。

Presentation API は、共有キャンバス (Shared Canvas) という考え方に基づきつつ、World Wide Web コンソーシアム (W3C) が規定する Linked Data の記述方法である JSON-LD、注釈の記述方法である Web annotation に従って、画像と注釈をとりまとめて一つの資料として記述するルールである。よく用いられる例として、写本の場合には、一つのページが一つの共有キャンバスであり、そこに画像や注釈などが URI で指定されて紐付けられる。そして、複数の共有キャンバスをまとめて順番付けしたものが一つの manifest ファイルと呼ばれ、一つの資料を指し示すことになる。manifest ファイルには、資料のメタデータやページを読んでいく方向等も記述できるようになっており、多様なメタデータも同時に共有でき、また、日本の古典籍に多い縦書きの資料の場合にも対応可能となっている。すなわち、この manifest の URI があれば、外部のシステムからであっても、一つの資料の情報から各画像、さらにはその注釈へと、IIIF のルールに従って辿り、指定し、取り出すことができるようになる。これこそがまさに、サイロに閉じ込められた画像を効果的・効率的に共有し相互運用できるようにするための仕組みとなっているのである。

作成された Presentation API のデータを活用する方法については、まず、Web での IIIF 対応画像ビューワということになるだろう。これには、筆者が確認できた限りでは、Web ブラウザ上で使えるフリーのものが少なくとも 5 種類公開されており、それらから分岐したバージョンも複数存在する。フリーのビューワの中でも特に注目されるのは、ウェルカム図書館、英国図書館等が開発に関わっている Universal Viewer と、スタンフォード大学、ハーバード大学等が開発の中心となっている Mirador である。いずれも、画像の拡大縮小等の基本的な機能は OpenSeadragon というフリーソフトの高精細画像ビューワを取り込むことで実現している。

Universal Viewer の大きな特徴は、音声や動画、3D 画像など、様々なタイプのメディアを表示できるようになっているという点である。また、画像に付与した翻刻テキストを検索することもできるようになっている。そして、画像をダウンロードする機能や「現在見ている画面」の URL を取得してそれを共有することもできる。一連の機能は Universal Viewer のデモサイトで確認できるので、興味がある方はぜひご参照されたい⁵。

一方、Mirador の大きな特徴は、IIIF 対応で公開されている各地のデジタル画像を複数並列表示して拡大縮小等ができることと、それらの表示画像にそれぞれビューワ上で注釈をつけられることである。たとえば、図 3 のように、各地で公開されている同じテキストを並べて比較するといったことが容易に可能なのである。ここから、各画面でそれぞれの画像を一つの資料としてペ

⁵ <http://universalviewer.io/examples/>

ージをめくっていったりできるようになっており、その点は、まさに IIIF の特徴を活かした機能であると言える。

注釈機能に関しては、IIIF で定められている注釈の記法に基づくデータを図 4 のように表示することができ、さらに、ユーザが注釈をつける機能も用意されている。このサイトでは、さらに、注釈中のキーワードをクリックするとそのキーワードで画像検索が行われるようになっている。なお、図 4 で表示されている注釈は、元々は IIIF とは異なるフォーマットで付加されたものだが、それを IIIF Presentation API に準拠する形に変換したことで、Mirador 上での表示が可能となった。そして、標準的な形式なので、他の IIIF 対応サイト・システムで読み込んで表示することも可能となっている⁶。

IIIF が示した解決策

さて、先述の三つの問題点に戻ってみると、デジタル画像がサイロに閉じ込められて相互運用性がないという問題は、IIIF の各種 API を通じて技術的には解決できていると言っていいだろう。アプリケーションソフトがリポジトリ毎に別々に作成されるために開発と維持に費用がかかるといふ点は、上記のようにいくつかの機関がフリーソフトウェアを共同で開発して皆で共有するという流れがすでに出来上がっており、開発・維持費用の問題を大いに改善している。そのようなことを可能にしたのは、画像配信に関する規格を IIIF で共通化したからに他ならない。また、ユーザが不統一な多くのインターフェイスに対応しなければならない問題については、Mirador や Universal Viewer など、どれか一つ、自分の使いたいビューワを選べば、どこの画像でもそのビューワで表示できるという環境が技術的には提供されるに至っている。三つの問題点は、技術的にはすでに解決されていると言っていいだろう。しかしながら、こういった取組みは、最終的には多くの機関がこれに対応してデファクト標準化してくれないことには実質的な有用性を得られない。次に、IIIF がどの程度の広がりを見せているのか、について見てみよう。

IIIF の広がり

2015 年 6 月、IIIF コンソーシアムが設立された。複数の研究機関によるプロジェクトとして始まった IIIF への取組みは、このときすでに大きな広がりを見せており、年会費 1 万ドルを要する IIIF コンソーシアムの設立にさえ 11 の研究図書館等が参加している。その後、IIIF コンソーシアムの加盟機関は 55 を数え、日本からも東京大学、京都大学、国立情報学研究所、国文学研究資料館、関西大学が参加している。IIIF の利用はコンソーシアムに加盟せずとも可能であることから、IIIF の広がりを見る数字としては、IIIF のディスカッショングループに参加している 568 名という数字が一つの参考になるだろう⁷。ここでは、IIIF の導入や利活用の手法、規格の改良など、IIIF に主体的に取り組む人たちが議論に参加している。

ここからは、コミュニティに参加する者としてのやや印象論な話になるが、IIIF は全体として、研究図書館のシステムエンジニアのコミュニティを軸としてボトムアップで広がっている傾向が

⁶ たとえば、神崎正英氏による Image Annotator⁶での SAT 大正蔵画像 DB の頁画像と注釈の表示を参照されたい。 <http://www.kanzaki.com/works/2016/pub/image-annotator>

⁷ <http://iiif.io/news/2016/11/15/newsletter/>

強いように思われる。IIIF がシステムエンジニア達に支持される理由には、その採用のしやすさという面もあるだろう。

まず、規格文書はオープンライセンスで公開されており、誰でも閲覧可能で再配布もできる。もちろん、規格の利用も無償で可能である。このことは単に使いやすさというだけでなく、ベンダーロックインの問題も回避することができ、採用に踏み切りやすい。

また、既に画像の公開システムが提供されている場合、それを排除することなく、共存する形で IIIF 対応サービスを提供することもできる。実際の例としては、たとえばフランス国立図書館の Gallica の場合には、専用ビューワで高精細画像の閲覧ができるようになっているが、同時に、ビューワで画像閲覧する際の URL を 2 箇所変更する⁸と IIIF 対応の URL になり、IIIF 対応ビューワでも同じ画像を読み込むことが可能となる。

さらに、現在 Web エンジニアの間で流行している JSON という記述フォーマットや、World Wide Web コンソーシアム(W3C)が近年進めている Linked Data の記述方法である JSON-LD、注釈の記述方法である Web annotation を採用していることなど、研究図書館のシステムエンジニアにとって取り組みやすいものとなっている点も、細かい話のようではあるが、やはり重要なポイントだろう。特に Web に関連するものの場合、流行している規格であれば、それに対応するためのツールが世界中で開発され、フリーで利用できるものも様々に用意されることになる。このことは、より身近な例としては、たとえば、我々が Unicode を扱おうとする場合に、かつては対応ツールが少なく Unicode 用の独自ツールの開発まで検討する場合もある等、かなり苦労したものだだったが、今ではほとんどのソフトウェアやシステムが Unicode に対応するようになったので文字コードをあまり意識せずに作業ができるようになりつつある、という状況に近いものがある。学習や開発、メンテナンス等に関してシステムエンジニアが気を配らなければならない要素が減ることは、特にシステムエンジニアを直接抱えていることの多い比較的規模の大きな欧米の研究図書館では、規格の採用を後押しするポイントになっているようである。

IIIF コンテンツを活用するための仕組み

IIIF がもたらした API 群は、オープンなライセンスとの組み合わせにより、Web 空間におけるデジタルコンテンツの共有に際し、基盤的提供機能の部分を切り分け、誰もが Web API を経由して外部からコンテンツの一部を直接指定してアクセスすることを可能とした。このことは、結果として、一次公開機関でなくてもコンテンツを様々な活用して再配布できるという状況を創り出した。より具体的に言えば、公開者でもなく閲覧者でもない第三者がコンテンツに即した有益かつ固有の情報を作成し、それをオープンに公開・共有できるようになったのである。そして、世界各地でこの API 群を活用できる利便性の高いツールが開発公開されるようになった。

比較的汎用性の高いツールについて見てみると、まず、トロント大学図書館が開発・公開している IIIF Toolkit with Mirador は、ジョージ・メイソン大学が開発・公開するメタデータ CMS、Omeka のプラグインとして作成されており、さらに、ヴァージニア大学図書館が開発・公開している時空間マッピング用 Omeka プラグイン、Neatline を組み合わせることによって IIIF 対応コン

⁸ URL 中の ark:/ の前に iiif/を挿入し、URL 末尾に/manifest.json と付記する。

テンツに対してユーザが簡単に時空間情報を付与することが可能となっている。各地で公開されている IIIF 対応コンテンツを対象にして別の Omeka サイトからアノテーションを付与することができるようになっており、この機能を通じて任意のデジタルコンテンツの中の任意の画像1枚に対して Google 検索が直接及ぶようにすることを誰にでも簡単にできるようにしてしまった。

別の方向性として、世界中の IIIF 対応コンテンツに含まれる任意の画像の任意の箇所を指定するという操作を繰り返した後、それぞれの箇所を順にたどってブラウジングすることを可能にした IIIF Curation Viewer も開発公開されている^[2]。

他にもいくつかの比較的汎用的に活用可能なツールが提供されているが、一方で、カスタムメイドされたシステムを通じた IIIF コンテンツの活用手法も様々なものが開発され、そのうちのいくつかは実運用に供されている。すでに比較的安定的に提供されているものを見てみると、SCTA⁹ は、スコラ哲学に関する研究教育サイトとして開発が続けられており、TEI 準拠¹⁰の異文を含むテキストを用いたテキストデータベースを基本としつつ、異文の情報は IIIF 準拠の頁画像を表示することによる確認が可能となっている。あるいはまた、後述するが、筆者らが構築し 2016 年 6 月に公開し、その後運用と改良を続けている SAT 大正蔵画像 DB¹¹は、6000 件以上の IIIF 準拠のアノテーションを含む仏教画像データベースであり、アノテーションを検索した上で IIIF 対応ビューワ Mirador を利用して複数画像を並列表示したり、それぞれの画像上にアノテーションとして各画像の属性をポップアップ表示したりできるようになっている。

IIIF の導入に伴う公開の在り方の変化

2016 年 11 月 1 日、笠間書院のブログに「文化資源のデジタルデータ流通に突きつけられた課題 — 国文学研究資料館のオープンデータ公開と永崎研宣氏による公開から考える [後藤 真 (国立歴史民俗博物館 研究部准教授)]」¹²という記事が公開された。近年広まりつつあるオープンデータ化の流れと IIIF に準拠した画像の公開が重なり合ったところに生じてくるデジタル資料提供機関の新たな課題について、筆者の活動を題材として論じたものである。事態の具体的な事柄や後藤氏の見解については上記のブログ記事を参照していただくとして、以下に、ここで指摘された課題を筆者なりに整理してみよう。

これまで、デジタル文化資料の公開と言えば、元になる資料を所蔵する機関・組織（以下、所蔵者）がデジタル化して直接に Web に公開するか（以下、所蔵者兼一次公開者）もしくは、そこと緊密に連携しデジタル資料公開を引き受けた機関・組織（以下、一次公開者）が閲覧のためのインターフェイスも含めたトータルな Web サービスを提供し、デジタル資料を公開する権利を独占しつつ、それに永続的な責任を持つという流れが主流だったように思われる。そして、再配布を明示的に禁じているケースも少なくなかった。しかし、昨今のオープンデータ化の流れは、デジタル資料の再配布を許可するというものであり、すなわち、当該資料のライセンスに従う限

⁹ <https://scta.info/>

¹⁰ <http://www.tei-c.org/>

¹¹ <https://dzkings.l.u-tokyo.ac.jp/SATi/images.php>

¹² http://kasamashoin.jp/2016/11/post_3796.html

り、誰でも公開者になることができるという状況になっているのである。このことは、デジタル文化資料の流通のある局面においては大きな変化をもたらしつつある。誰もが自分のところで公開できる資料は、様々に利活用されるだけでなく、一次公開者が何らかの事情で公開を停止した後もどこかで公開され利活用され続けることが可能となる。Internet Archive や Wikimedia 等、分野を問わず画像を公開・共有できるものもあれば、分野毎のポータルもあり、商用サービスも様々に存在しており、オープンデータ化すれば、必要だと思ってくれている人がいる限り、おそらくは、Web のどこかで生きながらえることができるだろう。また、結果として、一次公開者、あるいは、所蔵者兼一次公開者における一次公開者としての存在感はやや減じていくこともあるだろう。一方、所蔵者に関しては、デジタル代替物（元々の物理的な資料が存在し、それをデジタル化したもの）の場合、所蔵者はどこか、デジタル化したのはどこか、ということが資料の信頼性の決め手になることが少なくない。とりわけ来歴情報が重視される古典籍・美術品等においてはむしろ所蔵者の存在はますます強く意識されることになるだろう。同時に、そういったことが問題にならない資料については、所蔵者としての存在感もまた減じていくであろうことも視野に入れておく必要があるだろう。

これまでは「便利な Web サービス」を提供することが一次公開者の使命であり所蔵者がそれをも担うことは所蔵者の社会的使命の一環であるという意識があったように思う。しかし、オープンデータ化が様々な局面で「より便利な Web サービス」を生み出すことを期待するのだとしたら、同時にそれは、一次公開者としてどこまで Web サービスの利便性（特にインターフェイス）に注力するのか、ということのを再考しなければならないだろう。

IIIF が登場してきたのは、まさにこのような状況においてであり、オープンデータ時代のデジタル化資料一次公開者としての Web サービス提供の在り方を技術面から提示したものとなっている。すなわち、「コンピュータプログラムを用いて任意のサイトのデジタル化資料の任意の箇所を取り出して活用することができる上に、デジタル化資料を丸ごとダウンロードすることもできる」ことを実現する規格を提供したことで、一次公開者の手間を大幅に減らしつつ利活用性を高めることを可能としたのである。

一方、オープンなソフトウェア開発を論じた「伽藍とバザール」¹³が提起したように、この情報流通の仕方を集中と分散という観点からとらえてみると、オープンデータ化は分散志向との親和性が高い。一方、IIIF の場合、サイト間の垣根を超える手段を提供しつつも、現在普及している Web 技術の範囲では、注釈などの色々な付加情報を集約するにはむしろ一次公開者のサイトにデジタル化資料を置いておき、それを外部のサイトから利用することが想定されている。これはどちらかと言えば集中志向であると言っていいだろう。双方の志向にどう折り合いをつけていくか、ということも今後の課題になっていくだろう。

¹³ Eric S. Raymond, The Cathedral and the Bazaar, 1999,
<http://cruel.org/freeware/cathedral.html> (日本語訳)