

IIIF を利用した富士川文庫資料の再統合の試み

西岡 千文 (京都大学 附属図書館)

富士川文庫は、富士川游博士の旧蔵書である古医書のコレクションである。コレクションの資料は、京都大学、慶應義塾大学等の複数機関に分散されて所蔵されている。本稿では、京都大学図書館機構と慶應義塾大学メディアセンターによる「富士川文庫デジタル連携プロジェクト（通称：デジタル富士川）」について述べる。本プロジェクトの目的は以下の二点である。一点目は、富士川文庫資料を Web 空間で一堂に公開することで日本医学史の研究・発展に寄与することである。二点目は、両学が国際的な画像の相互利用促進の枠組みである IIIF (International Image Interoperability Framework) を利用して資料を公開していることから、IIIF の利点を活用した分散コレクションの仮想統合手法を検討・例示することである。プロジェクトの試行版サイトは2018年9月に公開された。本稿では、試行版サイト構築の過程で検討された文庫資料統合を実現する手法について述べる。最終的には、検索・ソート機能を有していることから DataTables を使用した手法が採用された。さらに、持続可能な統合サイトの運営に向けて、資料の追加・更新に伴う統合サイトの更新方法について、IIIF Change Discovery について言及しつつ、検討する。

Reunion of the Fujikawa Collection Powered by IIIF Chifumi Nishioka (Main Library, Kyoto University)

The Fujikawa collection is a collection of old medical records, which had been collected by Dr. Yu Fujikawa (1865-1940). Records in the collection are distributed among multiple institutions such as Kyoto University, Keio University, and Tokyo University. In this paper, we describe the “Digital Fujikawa -History of Medicine in Japan-”, which has been launched by Kyoto University Library Network and Keio University Media Center. This project aims to reintegrate the records in the Fujikawa collection on the web space. Since both universities make records public using IIIF (International Image Interoperability Framework), which is a framework for promoting mutual use of images, we will investigate the integration method using IIIF.

1. まえがき

富士川文庫は、富士川游博士 (1865~1940 年) の旧蔵書である古医書のコレクションである。コレクションの資料は、博士の畢生の大作である「日本医学史」の編纂のために収集されたといわれている。文庫を構成する資料は、京都大学、慶應義塾大学、東京大学等の複数機関に分散されて所蔵されている。京都大学附属図書館では、所蔵する富士川文庫資料の電子化に取り組んでおり、2018年4月26日に京都大学貴重資料デジタルアーカイブ¹で全点公開²を完了した[1]。慶應義塾大学メディアセンターでも資料の電子化が進展しており、慶應義塾大学デジタルコレクション³に

て順次公開している。コンテンツの国際展開を目指して、両学のデジタルアーカイブは、国際的な画像データ流通促進の枠組みである IIIF (International Image Interoperability Framework) [2]に準拠して資料を公開している。IIIF は、画像とそれに付随するデータ(メタデータ、注釈など)の相互利用を可能とするために、API を規定している。欧米の図書館・博物館・美術館に続いて、日本国内でも導入が広まっており、IIIF 準拠デジタルアーカイブを公開する機関は増加している。京都大学貴重資料デジタルアーカイブと慶應義塾大学デジタルコレクションは、ともに IIIF 準拠デジタルアーカイブである。

本稿では、京都大学図書館機構と慶應義塾大学メディアセンターの共同プロジェクトである「富士川文庫デジタル連携プロジェクト（通称：デジタル富士川）」について述べる。プロジェクトでは、IIIF を活用することで、散逸されて所蔵されている富士川文庫資料を Web 空間で再統合する。プロジェクトの目的として、下記の二点が挙げられる。

・両学が分散して所蔵する「富士川文庫」を一堂

1) 京都大学貴重資料デジタルアーカイブ。京都大学附属図書館が所蔵する医学分野の貴重資料コレクション「富士川文庫」全点の電子画像を公開しました。

<https://rmda.kulib.kyoto-u.ac.jp/collection/fujikawa>

2) 京都大学貴重資料デジタルアーカイブ。富士川文庫。
<https://rmda.kulib.kyoto-u.ac.jp/news/2018-04-26-0>

3) 慶應義塾大学メディアセンターデジタルコレクション。富士川文庫 (古医書コレクション)。

<http://dcollections.lib.keio.ac.jp/ja/koisho>

に公開することを通じて日本医史学の研究・発展に寄与する。

- IIF の利点を活用した分散コレクションの仮想統合手法について共同で検討・開発し、例示する。

2018年9月28日に、富士川文庫デジタル連携プロジェクト試行版サイト¹⁾（以下、試行版サイト）が公開された。試行版サイトには、両学が所蔵する富士川文庫資料計5,160点が公開されている。検索・ソートといった資料の探索に必要な基本的な機能が実装されている。試行版サイトでは、利用者は、二次利用を行うまで、資料の提供機関を意識せずに、資料を閲覧・探索できる。

本稿の構成は以下に示す。次章では、試行版サイト構築の過程で検討された文庫資料統合を実現する三手法を提示する。また、各手法の利点・課題を挙げる。第3章は、資料の追加・更新に伴う統合サイトの更新手法について検討し、第4章を結びとする。

2. IIF を活用した文庫資料の統合手法

両学のデジタルアーカイブは、IIF が公開する API のうち、IIF Image API と IIF Presentation API のみを利用している。IIF Image API は画像データの取り扱いについての規定を定める。IIF Presentation API は、ビューワでの画像の表示順序や資料のメタデータといった、資料を表示するために必要な情報の記述方法を定める。

本章では、IIF を利用した文庫資料の統合手法を示す。具体的には、試行版サイト開発の過程に検討された IIF Presentation API のコレクションオブジェクトを使用した統合手法、Mirador を使用した統合手法、DataTables を使用した統合手法について報告する。試行版サイトでは、最終的に DataTables を使用した統合手法が採用された。なお、試行版サイトの Web サーバは ftp のみでのアクセスが許容されている。よって、Omeka S を利用した統合手法[3]などソフトウェアのインストールを必要とする手法は検討していない。利点としては、ソフトウェアの更新などに対応する必要がないため、管理に要する手間が少ない点が挙げられるだろう。

再統合の対象となる資料数は、5,160 点（京都大学附属図書館所蔵分 4,710 点、慶應義塾大学メディアセンター所蔵分 450 点）である。将来、慶應義塾大学メディアセンターより 1,000 件以上の資料が順次追加される予定である。

1) 富士川文庫デジタル連携プロジェクト。
http://www.kulib.kyoto-u.ac.jp/rdl/digital_fujikawa/

【IIF Presentation API のコレクションオブジェクトを使用した統合手法】

IIF で規定されている仕様を活用した手法として、IIF Presentation API で定義されている `sc:Collection` オブジェクト（以下、コレクションオブジェクト²⁾）を使用した手法が挙げられる[4]。コレクションオブジェクトは、資料に相当する `sc:Manifest` オブジェクト（以下、マニフェスト）の URI ならびに表示するマニフェストのラベル（通常は資料タイトル）の JSON-LD 形式のリストである。作成したコレクションオブジェクトをコレクションファイルとして保存する。IIF コミュニティで利用・開発が盛んなビューワである Universal Viewer でコレクションファイルを読み込むことで、Universal Viewer の左部に表示される Index タブで、コレクションに含まれている資料の一覧が表示される（図 1）。一覧における資料の順序は、コレクションオブジェクトにおける資料の順序と一致する。また、コレクションオブジェクトのリストにおいて先頭に配置されている資料の画像が、ビューワに読み込まれて表示される。

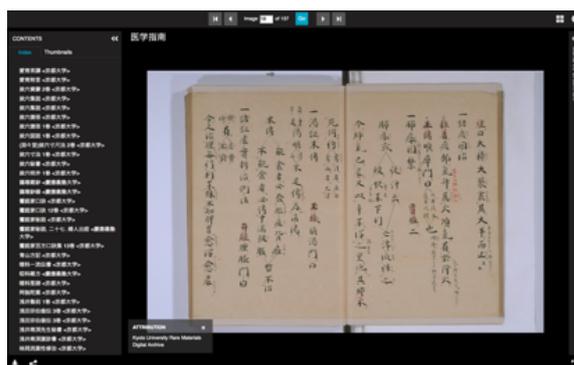


図 1 IIF Presentation API のコレクションオブジェクトを使用した統合手法

Figure 1 Reunion of the collection using `sc:Collection` specified in the IIF Presentation API.

この統合手法は、IIF で標準として定められている仕様に従うことから、後述の二手法と比較すると、最も IIF を活用しているといえる。しかし、コレクションオブジェクトに対応している主要なビューワが、2018年10月現在では、Universal Viewer のみである。Universal Viewer はコレクションオブジェクトに対応しているものの、コレクション内での検索・ソートといった機能は実装されていない。コレクションオブジェクトを利用した統合手法を採用するためには、より多くのビューワ

2) 本稿では以降、「コレクション」は IIF Presentation API で定義される `sc:Collection` オブジェクトを意味する。

一ワのコレクションオブジェクトへの対応や、コレクションオブジェクト内での検索・ソートといった機能の実装が必要となってくる。

【Mirador を使用した統合手法】

Mirador は Universal Viewer 同様、IIIF のコミュニティで最も利用されているビューワである。画像の比較や注釈・翻刻といった機能が充実していることから、研究者による利活用が期待されている。Mirador では、予め読み込む資料のリストをソースコードで指定できる¹⁾。資料画像が表示されていないときに Mirador 中央に表示される「アイテム追加」をクリックすると、指定された資料のリストがタイトル、所蔵機関とそのロゴ、さらにサムネイルとともに表示される (図 2)。当該画面は、資料画像が表示されているときでも、画像の右上に位置するアイコンをクリックし、「オブジェクトを追加」を選択することでも、表示することができる。また、上部には検索ボックスが提供されており、キーワードによって資料を絞り込むことが可能である。

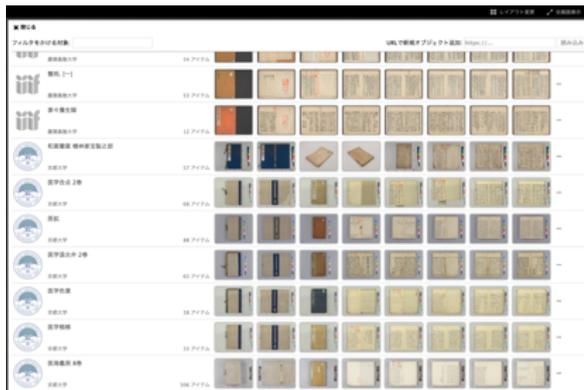


図 2 Mirador を使用した統合手法
Figure 2 Reunion of the collection using Mirador.

しかし、今回対象となる富士川文庫には、5,000 件以上の資料が含まれる。この場合、5,000 件以上のマニフェストファイルと、各資料のサムネイル画像 (表示画面サイズに依存するが各資料に対して約 5~9 枚) に対してアクセスが発生する。そのため、資料件数が大きい文庫の再統合手法としては不適切であると判断した。

1) 例として下記のサイトを挙げる。各スクリーンの左上に表示されている「新しいオブジェクト」というアイコンをクリックすると、図 2 にあるような資料のリストが表示される。サイトのソースコードの 27~75 行目にかけて、資料のマニフェストファイルの URL と所蔵機関のデータが格納されている。

<http://projectmirador.org/demo/>

【DataTables を使用した統合手法】

DataTables²⁾ は jQuery のプラグインであり、検索・ソートといった機能を HTML テーブルに追加する。本手法は、DataTables によって機能が追加された HTML テーブルに、資料を一覧表示する (図 3)。各行に資料タイトル (原題, カナ, ローマ字), 著者, 出版年 (西暦), 所蔵機関が表示される。これらのメタデータは JSON ファイルに格納されている。資料タイトル (原題) には、各機関における当該資料のページへのリンクが提供されている。当該資料のページに表示されている全メタデータが、マニフェストファイルに格納されているのであれば、ビューワで全メタデータを閲覧することが可能である。しかし、一部のメタデータ³⁾は機関が提供する当該資料のページでのみ提供されていることから、リンクを提供している。さらに、各行には IIIF アイコン (マニフェストファイルへのリンクが提供されているアイコン) が表示される。また、Universal Viewer, Mirador, IIIF Curation Viewer [5] で画像を表示するためのアイコンが提供されている。Universal Viewer と Mirador は試行版サイトにてビューワが構築されている。各機関から提供されているマニフェストファイルをそれぞれのビューワで読み込むことで、資料を表示する。試行版サイトでは、DataTables を使用することで、検索・ソートを可能としている。デフォルトでは、資料タイトル (カナ) 順で資料が一覧表示される。検索対象はアイコンを除く全カラムである。また、アイコンを除いた各カラムでソート可能である。

資料タイトルとして、京都大学附属図書館は新字体、慶應義塾大学メディアセンターは旧字体を主に使用している。DataTables で提供されている検索機能では、新字体・旧字体の統合検索が不可能である。例えば、「医学」と検索した際に「醫學」はヒットしない。しかし、慶應義塾大学メディアセンター所蔵資料には、新字体を使用したタイトルが別タイトルとしてメタデータに格納されている。別タイトルを隠しカラムで提供することで、新字体のクエリで慶應義塾大学メディアセンター所蔵資料を検索することを可能とした。しかし、旧字体のクエリでは、慶應義塾大学メディ

2) DataTables. <https://datatables.net/>

3) 例として、下記のページで提供されている冊数・注記といった情報は、統合サイトやマニフェストファイルからは確認することができない。

<https://rmda.kulib.kyoto-u.ac.jp/item/rb00000767>

Search: 医学

書誌タイトル	書誌タイトル (カナ)	書誌タイトル (ローマ字)	著者	出版年	所蔵
医学色葉	イガク イロハ	Igaku iroha	石心子著		京都
医学温古井 2巻	イガク オンコベン	Igaku onkoban	山口実斎著	1743	京都
医学楷梯	イガク カイテイ	Igaku kaitei	向田迪(応斎)著		京都
醫學管鑑. 總目	イガク カンスイ	Igaku kansui	山田, 業廣		慶應
醫學管鑑. 内集. 十三	イガク カンスイ	Igaku kansui	山田, 業廣		慶應
醫學管鑑. 内集. 十五	イガク カンスイ	Igaku kansui	山田, 業廣		慶應
醫學管鑑. 内集. 一	イガク カンスイ	Igaku kansui	山田, 業廣		慶應
醫學管鑑. 内集. 二	イガク カンスイ	Igaku kansui	山田, 業廣		慶應
醫學管鑑. 内集. 三	イガク カンスイ	Igaku kansui	山田, 業廣		慶應
醫學管鑑. 内集. 四	イガク カンスイ	Igaku kansui	山田, 業廣		慶應

Showing 1 to 10 of 145 entries (filtered from 5,160 total entries)

Page 1 of 15

※ 所蔵の表記について：「京都」＝「京都大学附属図書館」「慶應」＝「慶應義塾大学信濃町メディアセンター」（二次利用条件は所蔵機関によって異なるため、要確認）

図3 DataTablesを使用した統合手法（試行版サイトに採用）

Figure 3 Reunion of the collection using DataTables. This implementation has been released as a trial site of the project.

アセンター所蔵資料のみがヒットし、京都大学附属図書館資料を検索することはできない。今後の改善が必要である。

検索・ソートといった資料の探索に不可欠な機能が提供されていることから、DataTablesを使用した統合手法を試行版として公開した。統合サイトでは、ユーザは資料の提供機関を意識せずに資料を探索・閲覧することが可能だ。しかし、二次利用条件は所蔵機関・資料によって異なるので、明瞭な条件表示が必要となる。特に、ビューワではマニフェストファイルに記載されているメタデータのみが使用されるので、マニフェストファイルに条件表示を明記することが必須となる。

3. コンテンツの追加・更新手法

京都大学附属図書館所蔵分の富士川文庫資料は全点電子化・公開が完了しているが、慶應義塾大学メディアセンターでは電子化作業が進行しており、追加公開が予定されている。また、IIIFの利用の広まりとともに、本稿で示したような複数機関による資料の連携が、分野横断ポータルへの公開などをはじめとして、多数期待される。

持続可能な統合サイトには、機関のデジタルアーカイブに資料が追加・更新された際に、それらを反映させる手法が必要となる。3.1節では、試行版サイト構築時に採用された資料の登録手法

とともに、現在採用されている資料の追加・更新手法について述べる。3.2節では、今後の資料の追加・更新手法をIIIF Change Discovery APIを中心に検討する。

3.1. 試行版サイトにおける資料の追加・更新手法

【試行版サイト構築時における資料の登録手法】
試行版サイトでは、資料タイトル(原題, カナ, ローマ字), 著者, 出版年(西暦)といった基本的なメタデータが表示される。これらのメタデータと資料のマニフェストファイルのURIを収集することが必要となる。試行版サイト構築時には、機関に応じて異なる登録方法が採用された。

慶應義塾大学メディアセンター所蔵資料については、登録する資料のマニフェストファイルのURI一覧を提供してもらい、それらのURIにアクセスしてメタデータを取得する。それに対して、京都大学附属図書館所蔵資料については、資料のマニフェストURIとメタデータ一覧は、エクセルファイルによって提供された。エクセルファイルでの提供が選択された理由として二点挙げられる。一点目は、京都大学附属図書館所蔵資料のマニフェストファイルには、一部のメタデータ(資料タイトル(カナ, ローマ字)など)が含まれていないことである。二点目の理由としては、

京都大学附属図書館所蔵の資料件数が大きい(約5,000件)ことが挙げられる。資料件数が大きい場合、数秒ごとに個別の資料のマニフェストファイルにアクセスしてメタデータを取得すると、処理に要する時間は長くなる。それよりも、CSVファイルやエクセルファイルなどでメタデータを提供してもらい、それらを一括登録したほうが、要する時間は圧倒的に短い。

IIIF(第3.2節で詳述)国際標準の観点からは、マニフェストファイルを利用した登録が望ましい。資料件数が大きい場合でも短時間で一括登録できるように、マニフェストファイルの集合をバルクで提供するといったことも検討していく必要がある。

【試行版サイトにおけるコンテンツの追加・更新手法】

資料が更新・追加された際には、これらの資料のメタデータを取得して統合サイトに反映させる必要がある。京都大学附属図書館所蔵資料については、既に全点の電子化・公開が完了しているため、追加・更新がなされる可能性は低く、追加・更新方法については検討されていない。慶應義塾大学メディアセンター所蔵資料については、資料が追加された際に追加された資料のマニフェストファイルのURIのリストを提供してもらう。構築時における資料の登録と同様に、それらのURIにアクセスしてメタデータを取得し、試行版サイトに反映させる。

3.2. 今後に向けた資料の追加・更新手法

2018年8月現在、IIIFではIIIF Change Discovery API[6]の開発が進行している。IIIF Change Discovery APIは、IIIFで公開されている資料の追加・更新・削除に関する情報の提供方法・取得方法について規定しており、統合サイトなどでの活用が想定されている。コンテンツ提供側はW3C Activity Streams¹⁾の仕様に従って、資料の変更を記述してシリアルライズする。コンテンツ提供側とコンテンツ利用側(統合サイトなど)のインタラクションは、ResourceSyncフレームワーク²⁾に基づいてなされる。

以上のように、IIIF Change Discovery APIはコンテンツの同期方法を定める。しかし、その他のIIIFのAPI同様、機械可読な資料の書誌情報の記

述に関する標準は定めない³⁾。IIIF Presentation APIは、各用途に沿った形式の機械可読なメタデータは別ファイルで提供し、そのファイルのURIをマニフェストのプロパティ seeAlso に記載することを定める。IIIF Change Discovery APIでは、マニフェストのみならず、seeAlsoで与えられた書誌情報のファイルも必要に応じてクロウリングし、インデキシング等に利用することを想定している。しかし、調査[7]によると、seeAlsoを利用している機関は僅かである。今後、seeAlsoの利用がどの程度広まるかについても不透明である。さらに、seeAlsoが利用されている場合でも、フォーマットやメタデータスキーマは各機関で異なる[7]。よって、資料の統合の際には、関係機関による調整が必要となる。

4. あとがき

本稿では、京都大学図書館機構と慶應義塾大学メディアセンターによる共同プロジェクトである「富士川文庫デジタル連携プロジェクト」とともに、IIIFを利用した文庫資料の統合手法を示した。IIIF Presentation APIは、資料を文庫として表示するために、コレクションを用意しているが、対応しているソフトウェアは限られている。また、コレクションにおける検索・ソート等の機能は提供されていない。よって、DataTablesを利用した手法が試行版として公開された。DataTablesを使用することで、キーワードによる資料の検索、資料のソートが可能となる。本稿ではさらに、資料の追加・更新の際に自動的に統合サイトへ反映させる手法として、IIIF Change Discovery APIの使用について言及した。IIIFは、資料の書誌情報の記述に関する標準のフォーマットやスキーマについて定めていない。大規模な機関連携へと発展させていく際には、コミュニティでの合意の形成が必要不可欠となる。

富士川文庫デジタル連携プロジェクトでは、近い将来、慶應義塾大学メディアセンター所蔵の未公開資料の追加公開と連携機関の拡大を予定している。今後は、ユーザである研究者よりフィードバックを収集し、統合サイトの機能の検討・開発・実装を進め、試行フェーズから本稼働フェーズに移行することを目指している。

1) W3C Activity Streams.

<https://www.w3.org/TR/activitystreams-core/>

2) ResourceSync Framework Specification - Change Notification.

<http://www.openarchives.org/rs/notification/1.0.1/notification>

3) IIIF Presentation APIのマニフェストでは、フィールド label 等が必須項目として定められている。labelには、多くの場合において、資料のタイトルが入力されている。しかし実際は、タイトルとして表示したい文字列を入力するフィールドであるので、絵画作品などの場合では、作者名+タイトルが値として入力されていることもある。

謝辞

本稿で述べたプロジェクト「デジタル富士川」は、慶應義塾大学メディアセンターの保坂睦氏、入江伸氏、稲木竜氏、五十嵐健一氏、ならびに京都大学附属図書館の赤澤久弥氏、大村明美氏、富岡達治氏、北條風行氏、著者らを中心とした関係者らによって企画・遂行されている。著者は、試行版サイトの検討・実装を担当した。本稿執筆にあたっては、関係者各位から多大なサポートを受けたことに感謝する。

また、実装の検討にあたっては、2018年1月25日に開催された神崎正英氏による勉強会[4]（於：慶應義塾大学メディアセンター）から、IIIFを活用した文庫資料の統合方法に関するアドバイスをいただいた。ここに謝意を表す。

原稿を注意深く読んでくださり、有益なコメントをくださった匿名の査読者への感謝を表す。

京都大学附属図書館所蔵の富士川文庫の悉皆調査にあたっては、京都大学人文科学研究所・教授の武田時昌氏、森ノ宮医療大学・教授の長野仁氏に多大なるご協力をいただいたことに感謝する。

京都大学附属図書館が所蔵する富士川文庫の電子化は、下記の助成を受けて実施された。

- ・「京都大学附属図書館所蔵『富士川文庫』保存・公開のための修復・電子化事業 -わが国の医学の歴史を俯瞰する研究基盤構築のために-（機能強化経費）」
- ・国文学研究資料館「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」
- ・京都大学重点戦略アクションプラン オープンアクセス推進事業

慶應義塾大学メディアセンターが所蔵する富士川文庫資料の電子化は、下記の助成を受けて実施された。

- ・国文学研究資料館「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」
- ・慶應義塾大学メディアセンター本部／信濃町メディアセンター2006～2018年度事業計画（経常費）

参考文献

- [1] 西岡千文. 京都大学貴重資料デジタルアーカイブの紹介. 伝統鍼灸, 2018, Vol. 45, No. 1, p.36-42. 入手先 <<http://hdl.handle.net/2433/234685>> (参照 2018-10-30)
- [2] Snyderman, S., Robert S. and Tom C.: The International Image Interoperability Framework (IIIF): A community & technology approach for web-based images, *In Proceedings of the Archiving Conference*, pp.16-21, Society for Imaging Science and Technology (2015).
- [3] 中村覚, 永崎研宜. 日本国内の IIIF 準拠画

像に対する横断検索システムの構築. 研究報告 人文科学とコンピュータ (CH), 2018, No. 8, pp.1-6.

[4] 神崎正英. IIIF マニフェストの賞味法. 入手先

<<https://www.kanzaki.com/works/2018/pub/0125kda.html>> (参照 2018-10-30).

[5] 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. じんもんこん 2016 論文集, 2016, pp.117-124.

[6] IIIF Change Discovery API, 入手先 <<http://iiif.io/api/discovery/0.1/>> (参照 2018-08-20).

[7] IIIF Linking Properties Usage Survey, 入手先 <<https://github.com/benwbrum/fromthepage/wiki/IIIF-Linking-Properties-Usage-Survey>> (参照 2018-08-20)