IFLA News Media Section: Collecting, Transforming and Preserving News – for Research and the Public
April 27-28, 2017
Reykjavik, Iceland

Introduction to the International Image Interoperability Framework (IIIF), with a Focus on Digital Newspapers

Sheila Rabun
Community and Communications Officer, International Image Interoperability Framework (IIIF) Consortium
E-mail address: srabun@iiif.io

Abstract:

The International Image Interoperability Framework (IIIF) is a growing community of libraries, museums, software firms, and digital image repositories working to define, develop, cultivate, and document shared technologies that support interoperability for web-based image delivery. Digital images provide access to an increasing number of cultural heritage materials on the web, including newspapers and other media. However, many of these resources have been locked in to siloed, bespoke websites that are often challenging to maintain, with limited functionality for end users. IIIF provides a solution to this problem, utilizing linked data and shared application programming interface (API) specifications to enable enhanced functionality, data portability, and sustainability for digital image repositories. As a community-driven initiative, IIIF relies on discussion and input from individuals and institutions involved in digital image repositories. This presentation will give an introduction to IIIF, with a focus on examples and benefits for digital newspaper repositories, current community activities, and ways to get involved.

Keywords: interoperability, digital images, annotation, shared canvas, IIIF

Introduction

The International Image Interoperability Framework, or IIIF (pronounced triple-EYE-eff), has been a buzzword in the digital library community, but what is it and what does it mean? IIIF is best described as a community that develops shared application programming interfaces (APIs), implements them in software, and exposes interoperable digital image content on the web. IIIF exists because scholars and researchers increasingly rely on digital image repositories for access to images of important artifacts like manuscripts, newspapers, maps, and other fragile, rare, and heavily studied cultural materials. Libraries, archives, and museums across the globe have been increasingly making their collection materials available to the public digitally, but online assets have essentially been locked into repository-based silos, with incompatible systems and varying user interfaces that place a burden on hosting institutions

and limitations on functionality for end-users. Digital collection websites have been developed essentially in isolation, and significant resources are needed to migrate and overhaul these snowflake systems as time goes on. Working independently, many institutions have found digital image repositories challenging to maintain, with constantly evolving user needs and ever-changing technology systems as digital content is migrated from one bespoke application to another.

Working together in response to these challenges, institutions and individuals from across the cultural heritage community developed the IIIF specifications to provide a standard practice for making digital images available online, allowing for the transfer and sharing of image pixels, metadata, and annotations across repositories and systems. Adoption of the IIIF specifications can provide end-users with the ability to compare images from across multiple repositories and interact with them through deep pan and zoom, image manipulation (size, quality, rotation, etc.), the ability to tag and annotate, search within annotations, and easily share work with others. Several image servers and clients that support IIIF already exist, allowing institutions to easily mix and match technologies and maintain their repositories while avoiding costly system overhauls, migrations and re-designs.
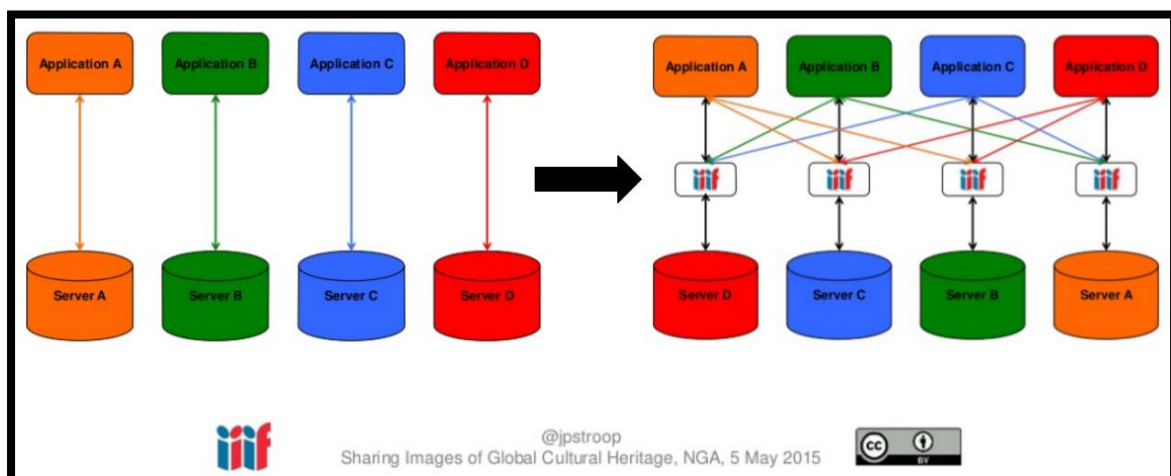

Figure 1. This figure depicts digital image repository silos on the left, and interoperability through IIIF APIs on the right. Image courtesy of Jon Stroop, Princeton University Libraries.

The IIIF community encompasses a large and growing group of interested and active individuals and organizations dedicated to leading and sustaining the IIIF. As a community-driven initiative, IIIF thrives on active discussion, input, and feedback from a wide array of diverse perspectives from libraries, museums, cultural heritage institutions, software firms, and other organizations working with digital images and audio/visual materials. Several interest-based community groups have formed within the community to address particular facets or applications of IIIF, including manuscripts, newspapers, software, and museums. There are also groups working on technical specifications for discovery of IIIF resources, and extending IIIF to apply to Audio/Visual materials as well. Any organization or individual interested in adopting the IIIF, developing software to support it, or giving feedback on the effort is welcome to get involved. The IIIF community currently encompasses over 100 organizations (Community), and the IIIF Consortium is comprised of 41 institutions that provide sustainability and steering for the initiative (IIIF Consortium).
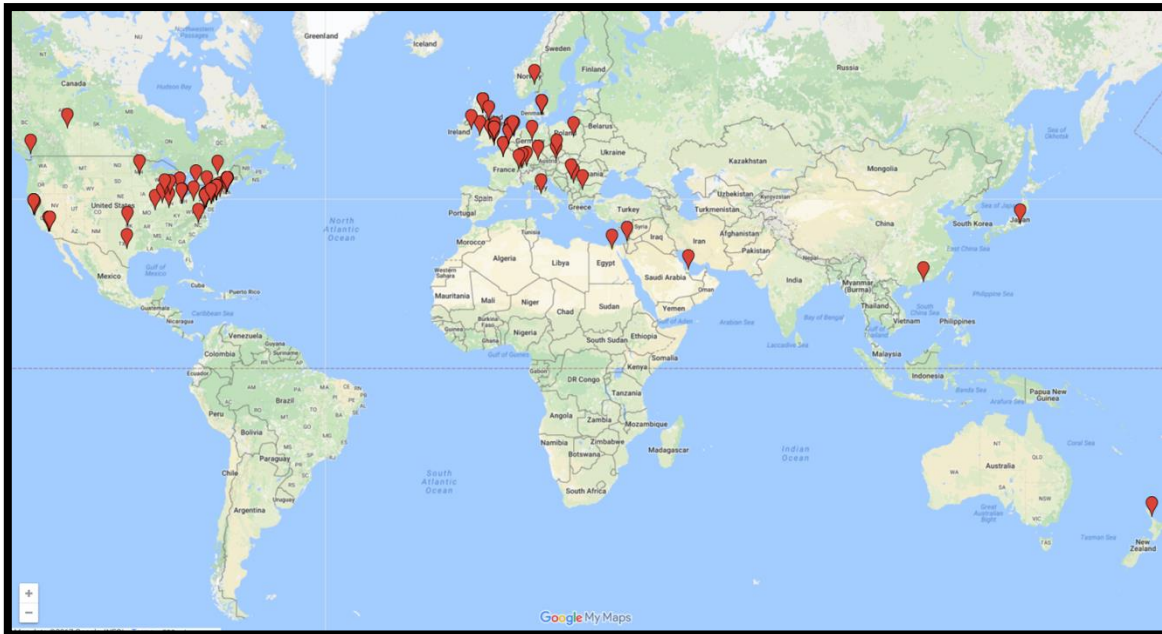
Figure 2. This map reflects the growing number of organizations adopting IIIF across the globe. Map courtesy of Tom Cramer, Stanford University Libraries.

Technical Details

IIIF is currently comprised of four specifications known as the IIIF Image API, IIIF Presentation API, IIIF Content Search API, and IIIF Authentication API, all based on the IIIF design patterns (Appleby, Crane, Sanderson, Stroop, & Warner, 2017). Adoption of the IIIF specifications provides benefits for both repositories and end users. Repositories can avoid vendor lock-in, complicated migrations, and expensive system overhauls by utilizing interoperable IIIF technologies. IIIF provides the ability to separate image delivery user interfaces from back end data stores, allowing repositories the ability to update image servers and databases without changing front end delivery, or vice versa, avoiding the need to re-architect the entire stack. With an active and growing community of organizations developing and supporting IIIF-compatible technologies, there are plenty of software options to choose from. A rich ecosystem of interoperable software, including many high quality, open source options, keeps licensing and operational costs low and predictable over time. Repositories can use any IIIF-compatible software for viewing, annotating, comparing, and manipulating images, regardless of the back-end server, and swap parts of the stack at any time, or run multiple components in parallel at once (IIIF FAQ). An initial list of clients and servers that support IIIF can be found at http://iiif.io/apps-demos/.

The IIIF specifications are based on existing web standards, using RESTful API construction, JSON-LD, and existing Web patterns, including the W3C Web Annotation standard, which simplifies processes for data migration and sharing. Repositories can publish images once and re-use often simply by changing HTTP(S) request parameters for an image's size, a region of an image, rotation, quality, and format. For example, rather than storing a master JP2, derivative JPG, and thumbnail for an image, one JP2 could be stored, using IIIF to manipulate and deliver different versions of the image on the fly. Thus, images held in one repository can be used in endless different ways both locally and in other sites and systems with no additional overhead. Institutions can combine content from across repositories both internally and with

external sources. This is especially helpful for related collections held at various different institutions, such as manuscripts that have been disbound and sold separately. As more and more museums and other institutions make digital images available via open access on the web, interoperability for research, exhibitions, and more becomes increasingly significant. However, for those images that may be restricted for whatever reason, IIIF allows for access control and can leverage existing SSO systems, via the IIIF Authentication API (IIIF Authentication API).

IIIF allows for advanced, interactive functionality for end users, including fast, rich, zoom and pan, manipulation of size, scale, region of interest, rotation, quality and format, the ability to crop, cite and share regions of an image, and create, read, and search for annotations. IIIF has native compatibility with the W3C Web Annotation Data Model, which supports annotating content on the Web (Web Annotation Working Group, 2017). Users can comment on, transcribe, and draw on image-based resources using the Web's inherent architecture and compare images from different repositories side by side with simple drag and drop.
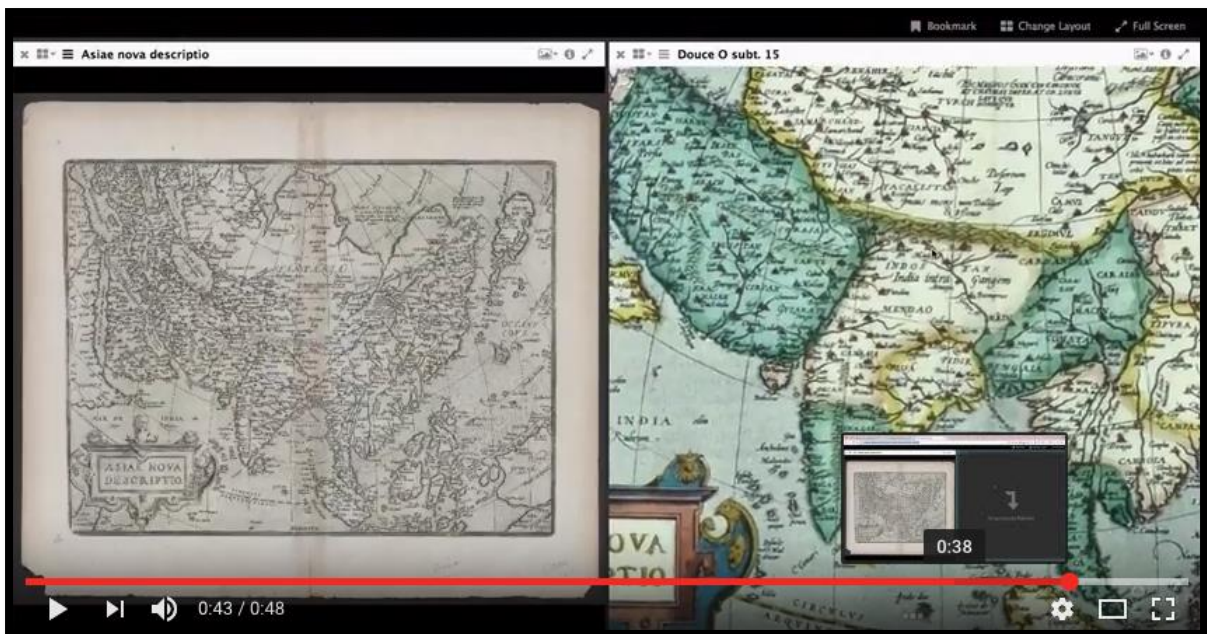

Figure 3. Screenshot from a recorded demo of the IIIF "drag and drop" functionality, featuring images from Stanford University Libraries (left) and the Oxford Digital Bodleian Libraries (right) as they were dragged and dropped into the Mirador Viewer (Stanford SUL, 2016).

The IIIF Image API provides instructions for the transfer of image pixels between various servers and clients across the web. The Image API provides the very basics of digital image interoperability, and there are different levels of compliance (Image API Compliance). The image API allows for manipulation of image pixels for changing size, rotation, quality, format, and highlighting regions of interest within an image. Clients can either request a full or partial image, or information about an image using the following standard image request URI syntax.

To request an image:

{scheme}://{server}{/prefix}/{identifier}/{region}/{size}/{rotation}/{quality}.{format}
http://www.example.org/image-service/abcd1234/full/full/0/default.jpg (IIIF Image API)

To request information about an image:

{scheme}://{server}{/prefix}/{identifier}/info.json
http://www.example.org/image-service/abcd1234/info.json (IIIF Image API)

The IIIF Presentation API takes interoperability further to allow for the transfer and sharing of information about an image, including descriptive metadata, structural metadata, rights information, related resources, and annotations. This way, users have access to more complete information to describe the images they are interacting with, especially title, table of contents and sequencing information if applicable, and any additional annotations related to the image. The IIIF Presentation API is based on the Shared Canvas data model, in which images and related information are essentially "painted" onto an abstract space known as a "canvas" via W3C web annotations and annotation lists. Typically a canvas consists of one page or one view of an object, although multiple canvases can be painted onto a single canvas, which is especially helpful for multi-spectral imaging or images missing content that can be reunified with the same image from a different resource.

The IIIF Presentation API leverages linked data, existing web architecture, and the Shared Canvas data model, to describe information about an object as represented in a JSON-LD document referred to as a IIIF manifest (IIIF Presentation API). Within the Shared Canvas data model, a manifest represents a physical object, like a book or a newspaper. Multiple manifests comprise a collection, and objects containing multiple pages or parts are comprised of sequences of canvases, each containing specific content, represented as annotations within the JSON-LD manifest. The IIIF Image and Presentation APIs thus work together to present to the user a digital surrogate of a physical object, with the ability to manipulate, annotate, and otherwise more fully explore cultural heritage objects on the web. The IIIF community is currently working on extending the IIIF Presentation API to include Audio/Visual materials, which will entail adding a time dimension to the Shared Canvas data model.
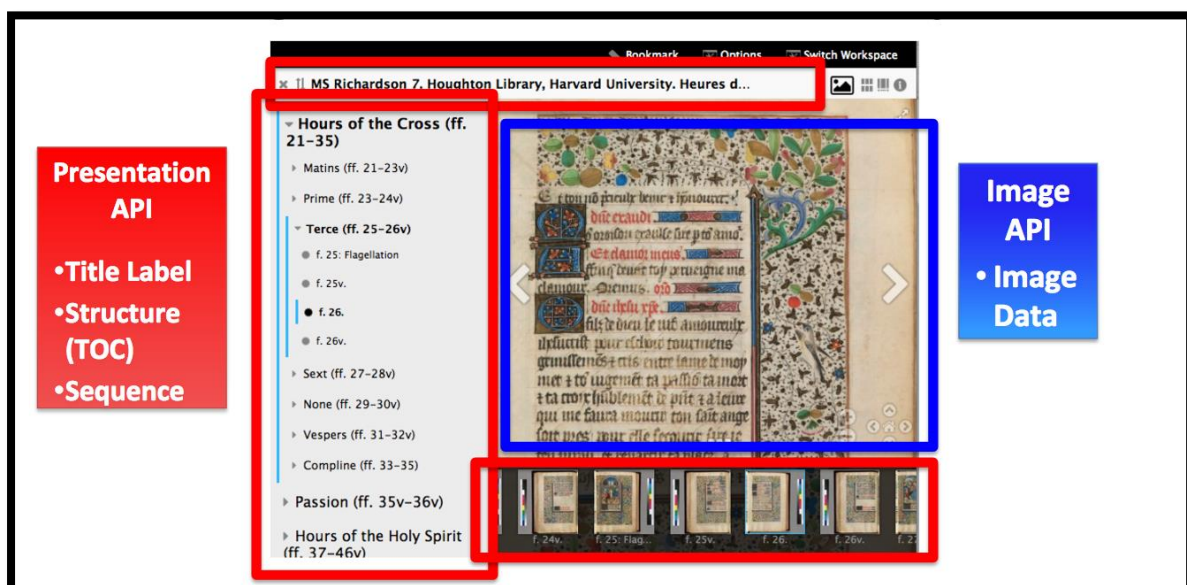

Figure 4. The services of the IIIF Image API (outlined in blue) and the IIIF Presentation API (outlined in red) combined form a full digital representation of a physical object such as a manuscript. Image courtesy of Rob Sanderson, The J. Paul Getty Trust.

The IIIF Content Search API allows for searching text within the annotation layer of a single IIIF resource, such as a manifest or collection, for example searching for a keyword within an issue or collection of issues of a newspaper (IIIF Content Search API). The IIIF Authentication API provides the option for repositories to make restricted content available through IIIF. According to the IIIF Auth spec, "Open access to content is desirable, but internal policies, legal regulations, business models, and other constraints can require users to authenticate and be authorized to interact with some resources. The authentication process could range from a simple restriction by IP address or a click-through agreement, to a multi-factor scheme with a secure identity provider" (IIIF Authentication API).

IIIF for Digital Newspapers

The IIIF Newspapers Community Group meets virtually once a month to share, discuss, and determine best practices for IIIF as it applies to digitized newspapers (IIIF Newspapers Community Group). Chaired by Karen Estlund of Pennsylvania State University Libraries and Glen Robson of the National Library of Wales, the group has been working to leverage and document the technical details and benefits of IIIF for digital newspaper collections, with considerations for the content aspects unique to newspapers, such as columns of text, use of OCR, publication date, and non-contiguous sections. Anyone with an interest in digital newspapers is welcome to join the IIIF Newspaper group calls and discussions.

The IIIF Newspapers group has mapped general newspaper metadata to the Shared Canvas data model used by IIIF:

Newspapers	IIIF
Title	Collection
Issue	Manifest
Edition	Manifest
Article	Range
Page	Canvas
Image	Image
Alto	Annotations

Figure 5. The IIIF Newspapers Community Group has mapped general newspaper metadata to the IIIF data model. Image courtesy of Karen Estlund, Pennsylvania State University Libraries.

Adoption of IIIF for digital newspapers is led by institutions including the National Library of Wales, North Carolina State University (NCSU) Libraries, and the Bavarian State Library. The National Library of Wales has made all of their newspapers available via the IIIF Image, Presentation, and Content Search APIs, using the Universal Viewer (see http://newspapers.library.wales/). North Carolina State University Libraries have also made their newspapers available via Image, Presentation, and Content Search (Ronallo & Rabun, 2016). The Bavarian State Library recently announced the release of roughly 350 newspapers now available through IIIF (see for example: https://digipress2.digitale-sammlungen.de/). The

Maryland Institute for Technology in the Humanities (MITH) have been exploring the use of the Mirador Viewer for digital newspapers (Tai, 2016). Mirador support IIIF, allowing users to view multiple images side by side at the same time from within or across repositories.

Penn State is currently working with the University of Nebraska, the University of Oregon, and others to launch the Open Online Newspaper Initiative (Open ONI), which supports the IIIF image API and uses the RAIS image server and OpenSeadragon client (Open ONI). Open ONI is a fork of the Library of Congress' chronam software used for the Chronicling America historic newspaper site at http://chroniclingamerica.loc.gov. For awardees from the U.S. National Digital Newspaper Program (NDNP), it's now easier than ever to make digital newspaper images available via IIIF. Ed Summers at MITH developed a shim to convert NDNP METS XML metadata into IIIF JSON-LD. For information about the NDNP shim and other helpful resources, please see the Newspapers section of the Awesome-IIIF list on GitHub at https://github.com/IIIF/awesome-iiif#newspapers.

Additional written introductory materials for understanding IIIF include "The International Image Interoperability Framework (IIIF): A community & technology approach for web-based images" (Snydman, Sanderson, and Cramer, 2015) and "An Introduction to IIIF" (Crane, 2017). To get involved with the IIIF community and learn more, all interested parties are invited to join the IIIF-Discuss Google Group email list.

Acknowledgments

References

Appleby, M., Crane, T., Sanderson, R., Stroop, J. and Warner, S. (2017). IIIF Design Patterns. The International Image Interoperability Framework. Retrieved 14 April 2017 from http://iiif.io/api/annex/notes/design_patterns/

Community. The International Image Interoperability Framework. Retrieved on 13 April 2017 from http://iiif.io/community/

Crane, T. (March 2017). An Intro to IIIF. *Digirati*. Retrieved from http://resources.digirati.com/iiif/an-introduction-to-iiif/

IIIF Authentication API. (2017). The International Image Interoperability Framework Consortium. Retrieved on 13 April 2017 from http://iiif.io/api/auth/

IIIF Consortium. The International Image Interoperability Framework. Retrieved on 13 April 2017 from http://iiif.io/community/consortium/

IIIF Content Search API. (2017). The International Image Interoperability Framework Consortium. Retrieved on 13 April 2017 from http://iiif.io/api/search/

IIIF FAQ Frequently Asked Questions. The International Image Interoperability Framework. Retrieved on 13 April 2017 from http://iiif.io/community/faq/

IIIF Image API. (2017) The International Image Interoperability Framework Consortium. Retrieved on 13 April 2017 from http://iiif.io/api/image/

IIIF Newspapers Community Group. The International Image Interoperability Framework. Retrieved on 13 April 2017 from http://iiif.io/community/groups/newspapers/

IIIF Presentation API. (2017). The International Image Interoperability Framework Consortium. Retrieved on 13 April 2017 from http://iiif.io/api/presentation/

Image API Compliance. The International Image Interoperability Framework. Retrieved on 13 April 2017 from http://iiif.io/api/image/2.1/compliance/

Open ONI. GitHub. Retrieved 14 April 2017 from https://github.com/open-oni/open-oni

Ronallo, J. and Rabun, S. (12 Sept. 2016). North Carolina State University (NCSU) Joins the IIIF Consortium, Announces Exciting IIIF Implementation. The International Image Interoperability Framework. Retrieved 14 April 2017 from http://iiif.io/news/2016/09/12/ncsu-joins-iiifc/

Snydman, S., Sanderson, R. and Cramer, T. (2015). The International Image Interoperability Framework (IIIF): A community & technology approach for web-based images. Archiving Conference, 16–21.https://stacks.stanford.edu/file/druid:df650pk4327/2015ARCHIVING_IIIF.pdf

Stanford SUL. (26 August 2016). IIIF drag'n'drop features in Stanford University Libraries online catalog. YouTube. Retrieved 13 April 2017 from https://youtu.be/uih5JuQnYuo

Tai, X. (15 August 2016). Final Report. Retrieved 13 April 2017 from http://taixiaoyu.com/potato/project/mith/mith-report.pdf

Web Annotation Working Group. (23 February 2017). Web Annotation Data Model. W3C Recommendation 23 February 2017. Retrieved 13 April 2017 from https://www.w3.org/TR/2017/REC-annotation-model-20170223/