IIIF & ResourceSync: Supporting discovery

Mark A. Matienzo, Stanford University Libraries
@anarchivist / https://orcid.org/0000-0003-3270-1306
DPLAFest — Chicago, Illinois — April 20, 2017
International Image Interoperability Framework

A community that develops Shared APIs implements them in Software and exposes interoperable Content

http://iiif.io/
IIIF Community

http://iiif.io/community

- **IIIF Consortium**
  - Currently 38 state/national libraries, universities, museums, tech firms
  - Provides sustainability and steering for the initiative

- **Wider community**
  - 80+ CH institutions, companies, and projects using IIIF standards
  - iiif-discuss list = 670+ members
  - IIIF Slack = 300+ members

- **Community & Technical Specification Groups**
Shared APIs

http://iiif.io/api/

- **Image API**
  - Transfer image pixels, regions, etc.
  - Image manipulation

- **Presentation API**
  - Presentation of an object (pixels + navigation and metadata)
  - Easily share and re-use, mix and match content
  - Annotate content

- **Search API**
  - Search annotations

- **Authentication API**
  - Provide interoperability for access-restricted content
Software Implementations

https://github.com/IIIF/awesome-iiif
IIIF Content

All kinds of image resources: artworks, photographs, manuscripts, newspapers

Investigating AV and 3D
“Discovery” in IIIF

Finding interoperable resources

Two main concerns:

● How can users find IIIF resources?
● How can users then get those resources into an environment where they can use them?
Scoping the problem

What resources can be discovered?

Types of resources in IIIF:

- Content (Image API)
- Description (Presentation API)

The Image API does not provide description of image content, just technical and rights metadata.

Discovery requires Description resources to provide information about Content resources.
Presentation API

A **Manifest** provides just enough metadata (descriptive, structural, etc.) to drive a viewer.

A **Collection** groups **Manifests** or other **Collections**.

http://iiif.io/api/presentation/2.1/
Community work

IIIF Discovery Technical Specification Group

iiif.io/community/groups/discovery/

IIIF Discovery TSG scope:

- Crawling and harvesting
- Content indexing
- Change notification
- Import to viewers
The Presentation API does not include *semantic* descriptions, but can reference them using **seeAlso**.

IIIF (including the Presentation API) has a *resource-centric* view of the web, not a *service-centric* view (cf Sitemaps/ResourceSync vs OAI-PMH).
Examples
Basic Sitemaps at NC State

- Example demonstrates use of Simple sitemaps without any extensions, including ResourceSync
- Intended to expand upon existing practice of publishing sitemaps from digital collections
Sample of NCSU Sitemaps

Sitemap entry for manifests

<url>
  <loc>https://d.lib.ncsu.edu/collections/catalog/bh1141pnc004/manifest</loc>
  <lastmod>2016-12-13T15:38:19Z</lastmod>
</url>

Sitemap entry for landing page

<url>
  <loc>https://d.lib.ncsu.edu/collections/catalog/bh1141pnc004</loc>
  <lastmod>2017-03-27T19:33:52Z</lastmod>
</url>

*Courtesy Jason Ronallo, North Carolina State University*
Prototyping at Europeana

Exploring Sitemaps and extensions for discovery of IIIF resources for harvesting

- Partnership with University College Dublin and National Library of Wales
- ResourceSync satisfied key needs identified within requirements
- ResourceSync accommodated additional metadata prototyped in an IIIF Sitemap Extension
- Follows several synchronization paradigms
Example of NLW Sitemap Entry

Uses Sitemaps and IIIF Extension

```xml
<url>
  <loc>http://newspapers.library.wales/view/3320640</loc>
  <iiif:Manifest xmlns:iiif="http://iiif.io/api/presentation/2/">
    http://dams.llgc.org.uk/iiif/newspaper/issue/3320640/manifest.json
  </iiif:Manifest>
  <dct:isPartOf>http://dams.llgc.org.uk/iiif/newspapers/3320639.json</dct:isPartOf>
  <lastmod>2014-11-08</lastmod>
  <changefreq>monthly</changefreq>
  <priority>0.8</priority>
</url>

Courtesy Nuno Freire, Europeana
Example of UCD Resource List Entry

Uses Sitemaps and ResourceSync and DCMES as Extensions

Example of UCD Resource List Entry

Courtesy Nuno Freire, Europeana
Sample of UCD Resource List

Uses Sitemaps, ResourceSync, and Sitemap Image Extension

 Courtesy John Howard, University College Dublin
Conclusions

**Strengths**

- ResourceSync addresses core requirements for exposing IIIF resources for harvesting
- Can build on publication of existing sitemaps easily
- Leverages *Many-to-One, Selective Synchronization*, and *Metadata Harvesting* paradigms
- Can adopt additional extensions to implement needed features
- Plenty of opportunity to contribute; *need more prototypes*

**Challenges**

- IIIF community’s needs for discovery are not necessarily what other sitemap consumers want (e.g. Google)
- Identifying the primary resource influences structure
- Unclear whether search engines support custom extensions, and what ranking impact would be
Thank You!

Mark A. Matienzo, Stanford University Libraries
@anarchivist / https://orcid.org/0000-0003-3270-1306
DPLAFest — Chicago, Illinois — April 20, 2017