Updates on DPLA and Hydra-In-A-Box

Mark A. Matienzo <mark@dp.la>
Digital Public Library of America

Mid-Atlantic Fedora Users Group
30 November 2015
What is DPLA?
DPLA is...

A *portal* for discovery

A *platform* to build on

A strong advocate for a *public option*

A network of *partnerships*
A Portal for Discovery

http://dp.la/

A Wealth of Knowledge
explore 8,007,019 items from libraries, archives, and museums

Search the Library

Apps
The DPLA is a platform. Developers make apps that use the library’s data in many different ways. Here are just a few.  App Library »

News
DPLA Brings National Attention to the Blue Earth County Historical Society
Oct 2
Exhibitions & Primary Source Sets

Early Chinese Immigration to the US
Primary Source Set
By Hillary Brady, Digital Public Library of America

The advent of the railroad brought about many changes to the United States, including an early wave of Chinese immigration to America. Chinese laborers were the backbone of the Transcontinental Railroad’s creation, and worked diligently in other difficult industrial jobs for low wages. During this period, they faced intense racial discrimination both socially and politically. This culminated in the Chinese Exclusion Act of 1882, which placed new obstacles in the way of Chinese and other Asian immigrants looking to come to America. In particular, the act placed limits and exclusions on laborers, not wealthy businessmen from China looking to immigrate. The following set of resources provides photos, documents, and oral histories that help tell the story of this early period of Chinese immigration to the United States.
A Platform to Build On

DPLA hack!

http://dp.la/apps/
A Strong Public Option
Partnerships

Service Hubs

Mountain West Digital Library
South Carolina Digital Library
Empire State Digital Network
Digital Commonwealth (MA)
NC Digital Heritage Center
The Portal to Texas History
Digital Library of Georgia
Minnesota Digital Library
Kentucky Digital Library
Indiana Memory
Missouri Hub

Content Hubs

Government Publishing Office (GPO)
National Archives & Records Admin.
Univ. of Illinois Urbana-Champaign
University of Southern California
Biodiversity Heritage Library
The New York Public Library
Harvard University Library
California Digital Library
Smithsonian Institution
University of Virginia
J. Paul Getty Trust
David Rumsey
HathiTrust
ArtStor
The DPLA Ingestion Process

1. Harvest Source
2. Original Record
3. Mapper
4. DPLA MAP Resource
5. Enrichments
6. Indexer
Krikri and Heiðrún

Enrichments
Mappers
Mapping DSL
Parsers
Harvesters
Resque
Redis
Marmotta
PostgreSQL
Solr

Krikri views
Blacklight views
Krikri controllers
Blacklight controllers
Krikri models
Blacklight models
Index service

DPLA MAP models
ActiveTriples
rdf-marmotta

Sample Mapping

```ruby
Krikri::Mapper.define(:esdn.mods, :parser => Krikri::ModsParser) do
  provider :class => DPLA::MAP::Agent do
    uri 'http://dp.la/api/contributor/esdn'
    label 'Empire State Digital Network'
  end

  dataProvider :class => DPLA::MAP::Agent do
    providedLabel record.field('mods:note').match_attribute(:type, 'ownership')
  end

  isShownAt :class => DPLA::MAP::WebResource do
    uri record.field('mods:location', 'mods:url')
    .match_attribute(:usage, 'primary display')
    .match_attribute(:access, 'object in context')
  end

  preview :class => DPLA::MAP::WebResource do
    uri record.field('mods:location', 'mods:url')
    .match_attribute(:access, 'preview')
  end

  originalRecord :class => DPLA::MAP::WebResource do
    uri record_uri
  end

  sourceResource :class => DPLA::MAP::SourceResource do
    alternative record.field('mods:titleInfo')
    .match_attribute(:type, 'alternative')
    .field('mods:title')
end
```

https://github.com/dpla/heidrun-mappings/blob/master/heidrun/esdn_mods.rb
Hydra-In-a-Box

• 30 month National Leadership Grant from IMLS

• Focuses on fostering a national digital platform through community-based repository infrastructure

• Leverages and contributes to Hydra through code & community
Project Goals

• Development of turnkey, Hydra-based application that leverages and improves on core components.

• Development/integration of metadata aggregation & enrichment tools

• Connect components with DPLA hubs, current Hydra partners, and prospective Hydra adopters

• Work toward a hosted service
## Project Team

<table>
<thead>
<tr>
<th>DPLA</th>
<th>DuraSpace</th>
<th>Stanford</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dan Cohen</td>
<td>Debra Hanken Kurtz</td>
<td>Tom Cramer</td>
</tr>
<tr>
<td>Mark Matienzo</td>
<td>Jonathan Markow</td>
<td>Hannah Frost</td>
</tr>
<tr>
<td>Mark Breedlove</td>
<td>Bill Branan</td>
<td>Mike Giarlo</td>
</tr>
<tr>
<td>Tom Johnson</td>
<td>Andrew Woods</td>
<td>Gary Geisler</td>
</tr>
<tr>
<td>Audrey Altman</td>
<td>Tim Donohue</td>
<td>Rob Sanderson</td>
</tr>
<tr>
<td>Kenny Whitebloom</td>
<td>Carol Minton Morris</td>
<td>Joe Atzberger</td>
</tr>
<tr>
<td>Rachel Frick</td>
<td></td>
<td>Chris Beer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Erin Fahy</td>
</tr>
</tbody>
</table>

... and the larger Hydra community
Project Timeline

- May 2015-November 2017 (30 months)
- Development: March 2016-Nov. 2017
- Service development and community engagement: throughout project
Design Process

Discovery
- Phase 1
  - Stakeholder goals
  - Interviews & focus groups
  - Community & landscape surveys
- Phase 2

Information Architecture
- Phase 1
  - User personas
  - Requirements prioritization
- Phase 2
  - Conceptual sitemaps
  - Wireframes

Visual Design
- Phase 1
  - Visual design mockups
  - Visual design style guide

Infrastructure & Technical Exploration
- Phase 1
  - Community engagement
  - Collaborative development on core gems
- Phase 2
  - Technical prototyping
- Phase 3
  - Gap analysis
  - Community feedback

Development
- Phase 1
  - Formal development & user testing

Timeline:
- Discovery: Phase 1
  - August 2015
- Discovery: Phase 2
  - September 2015
- Info Arch: Phase 1
  - October 2015
- Info Arch: Phase 2
  - November 2015
- Tech Exploration: Phase 1
  - December 2015
- Tech Exploration: Phase 2
  - January 2016
- Tech Exploration: Phase 3
  - February 2016
- Final: March 2016
Design Process Discoveries

• 75% respondents have some digital content not managed in a repository

• Satisfaction levels
  • Users of hosted services tend to be more satisfied than users of local deployments

• Strengths and weaknesses of existing repository options

• System migration
  • 53% plan to migrate; most to a Fedora-based solution
  • Rest are “not sure” what’s next
Participation

• Contribute ideas for functionality: http://bit.ly/hydrainabox-ideas

• Agree to participate in an interview or focus group

• Help us test and evaluate the software as we release it

• Participate in software development
Stay Informed

• Visit our website and blog: [http://hydrainabox.org/](http://hydrainabox.org/)

• Public information list: [hybox-info@googlegroups.com](mailto:hybox-info@googlegroups.com)

• Contact us: [hybox-contact@googlegroups.com](mailto:hybox-contact@googlegroups.com)
Thank You!

Mark A. Matienzo <mark@dp.la>
Digital Public Library of America