A Platform for Research

Rachel L. Frick
Director of Business Development, Digital Public Library of America
rachel@dp.la
DPLA: A brief history

- October 2010
  - planning process began

- December 2010
  - Berkman Center for Internet & Society
  - Historical Materials
    https://dp.la/info/about/history/materials/

- 2012
  - Private 501C3 organization
  - Staff of 4
  - led by Executive Director Dan Cohen
A Wealth of Knowledge
explore 8,961,168 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 »

- Metadata Explorer
  Dean Farrell
  Metadata Explorer is an attempt to explore national digital libraries, pan-national digital libraries and libraries in general, in a different, hopefully useful way.

- DPLA Visual Search Prototype
  Schuyler Lindberg
  A prototype visual search interface that explores content from the Digital Public Library of America.

News
DPLAfest 2015: Come hack with us!
Apr 1
DPLAfest 2015: Dive into the DPLA Hub network
Mar 31
All News »

@dpla on Twitter

more than a search box
DPLA is a Network
**Metadata** from content and service hubs is harvested by Digital Public Library of America.

**Link** to digital surrogate of original object on content hub or service hub is made available through DPLA.
DPLA is Partnerships
Public libraries, 24%

Private collections, 1%

Museums, 11%

International, 2%

K-12 schools, 1%

Historical societies, 10%

Other, 2%

Society/foundation archives/libraries, 4%

State agencies, 2%

State libraries, 1%

State/national parks, 1%

University departments, 1%

University libraries, 22%

Historic properties, 2%

Federal libraries, 1%

County and municipal agencies, 2%

Corporate libraries/archives, 2%

Archives, 6%

Community colleges, 3%
DPLA amplifies partner collections
DPLA is a Platform
Endless possibilities

JSON-LD out

Metadata in
For Developers

API Codex

Welcome to the DPLA API Codex! This is where you can find authoritative documentation for the DPLA API and resources you can use to make the most of it. We're happy to have you. (Wondering why we built an API in the first place? Check out our philosophy.)

What now?

- If you've never played (or worked) with an API before, head on over to API Basics. We'll get you started accessing some of our data.
- If you're already familiar with APIs in general and you just want to know how to get at our data, skip the basics and dive right into requests (Make sure you get an API key first.)
- Once you've got some shiny new JSON-LD, how can you tell what it all means? In responses, we explore the structure of objects and define the fields within those objects that the API will return.
- Ready to start building something against our API, but want to know how we operate? We've got you. Check out our policies.
- Uh oh. You're lost. We're lost. Everything's broken. I'm confused. Time to check out our troubleshooting and FAQ page!
- Confused by the terminology we're using? Check out the glossary of terms.
- Did we miss/break/overlook/forget something? Tell us.

The essentials

Need to start right now? Here's a cheat sheet.

- http://api.dp.la/v2 is the base URL of the DPLA API.
- items and collections are the two resource types you can request.
- In order to pass requests to the API, you'll need to request an API key first. You also need to append &api_key=$YOUR_KEY (where $YOUR_KEY is your 32-character key) to any requests you make.
- All records are returned as structured, wonderful, JSON-LD objects. And we have reasons for that.
- We've got a boatload of fields you can query, and a whole bunch of ways you can query them.
- Dive head first into a bucket full of kitten metadata: http://api.dp.la/v2/items?q=kittens&api_key=. (Copy and paste the URL into your browser and add your API key on the end.)

Edit this entry.

http://dp.la/info/developers/
The Many Minds Principle: The best thing to do with your data, will be thought of by someone else
More minds, more insights

“Knowledge is power, openness is empowerment”

Dr. Rufus Pollack
Founder, Open Knowledge
Associate, CIPIL, University of Cambridge

Making an Open Information Age, Open Repositories, 2016
Culture Hack: Libraries & Museums Open for Making
The DPLA is a platform that enables new and transformative uses of our digitized cultural heritage. The DPLA’s application programming interface (API) and open data can be used by software developers, researchers, and others to create novel environments for learning, tools for discovery, and engaging apps.

**Metadata Explorer**
Dean Farrell

Metadata Explorer is an attempt to explore national digital libraries, pan-national digital libraries and libraries in general, in a different, hopefully useful way.

[App Home Page »](#)

**DPLA Item Flow**
Dean Farrell

A visualization showing the flow of hubs into the DPLA.

[App Home Page »](#)

**DPLA Partner Flow**
Dean Farrell

A visualization showing the flow of items into DPLA hubs from their content providers.

[App Home Page »](#)

**DPLA Depositors Visualization**
Dean Farrell

Hover over an institution’s wedge to see how many items it has contributed to DPLA.

[App Home Page »](#)

**DPLA Bookshelf Language Visualization**
Dean Farrell

Visually browse books by language and view them on the DPLA Bookshelf.

[App Home Page »](#)

**Term Frequency Map**
Dean Farrell

See how your search term(s) frequency changes through time in DPLA.

[App Home Page »](#)

**DPLA Licenses**
Dean Farrell

A treemap graphic that shows the top 575 licenses in the DPLA corpus, divided into four clusters.

[App Home Page »](#)

**DPLA Visual Search Prototype**
Schuyler Lindberg

A prototype visual search interface that explores content from the Digital Public Library of America.

[App Home Page »](#)

**WikipeDPLA**
Eric Phetteplace

See relevant results from the Digital Public Library of America on any Wikipedia article using this Google Chrome extension.

[App Home Page »](#)
Enabling (re)use
Over 100,000 Unique Rights Statements

Visualization by Dean Farrell
Current DPLA Rights Statements

- Copyright or "All Rights Reserved": 45%
- Other: 5%
- Government Documents: 7%
- Missing Rights Statement: 7%
- "Digital image copyright": 8%
- Public Domain: 25%
- Creative Commons: 7%
Moose Creek (Idaho)

<table>
<thead>
<tr>
<th>Creator</th>
<th>Shattuck, C. H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Created Date</td>
<td>1904-1910</td>
</tr>
<tr>
<td>Partner</td>
<td>Mountain West Digital Library</td>
</tr>
<tr>
<td>Contributing Institution</td>
<td>University of Idaho Library</td>
</tr>
</tbody>
</table>

**Description**
Mature Cedar, Moose Creek, Idaho. Some photographs appear in a report from C. H. Shattuck, University of Idaho, Dept. of Forestry to Major F.A. Fenn, Supervisor, Clearwater National Forest (Day-NW SD428.C55S5) and in an article by C. H. Shattuck at the UI Agricultural Experiment Station Bulletin no. 91, Dec. 1916 (Idaho S 53 E416). Jp2 grayscale, enhanced image displayed. Scanned with Epson Expression 1640XL scanner. Master file is a 1000 dpi...

**Location**
United States--Idaho

**Type**
image

**Subject**
Nezperce National Forest (Idaho)

**Rights**
Digital image copyright 2010, The University of Idaho. All rights reserved. For more information contact Special Collections and Archives, University of Idaho Library, Moscow, ID 83844-2350; http://www.lib.uidaho.edu/special-collections/.

**URL**
Europeana Rights Over Time

Image credit: Kennisland (CC-BY-SA 3.0)
March 2014, Knight Foundation Grant

DPLA, Europeana, Creative Commons, & Kennisland

Develop a high level, standardized, international rights statements in a neutral namespace, expressed as a linked data controlled vocabulary
WORKING GROUP LEADERS

- Paul Keller (co-chair) - Kennisland/Europeana
- Emily Gore (co-chair) - DPLA

Subgroups

- Mark Matienzo (technical co-chair) - DPLA
- Antoine Issac (technical co-chair) - Europeana
- Dave Hansen (statements co-chair) - UNC/DPLA
- Julia Fallon (statements co-chair) - Europeana

WORKING GROUP MEMBERS

- Greg Cram - NYPL/DPLA
- Melissa Levine - UMich/DPLA
- Tom Johnson - DPLA
- Amy Rudersdorf - DPLA
- Karen Estlund - PennState/DPLA
- Richard Urban - FSU/DPLA
- Maarten Zeinstra - Kennisland/Europeana
- Patrick Peiffer - BNL/Europeana
- Lucie Guibault - Institute Info Law/Europeana
- Marie-Claire Dangerfield - Europeana
- Diane Peters - Creative Commons
- Matt Lee - Creative Commons
DPLA as Research
Collections as Data

http://www.stanford.edu/group/ruralwest/cgi-bin/drupal/visualizations/us_newspapers
Mining Cultural Heritage

What?

In research projects:

- Basic text mining: e.g. Word Clouds
- Network analysis
- Topic Modelling

Text and Data Mining

Using Cultural Heritage Data: Opportunities and challenges
Related research efforts

  
  https://www.ideals.illinois.edu/bitstream/handle/2142/145/shreeves05.pdf?sequence=2

• “Beyond Size and Search” Carole L. Palmer, Oksana L. Zavalina, and Katrina Fenlon
  
  https://www.ideals.illinois.edu/bitstream/handle/2142/18655/BeyondSizeAndSearch_final.pdf?sequence=2

  
DPLA Data Analysis

• Subject Descriptions in DPLA
  – Mark Phillips, UNT
  • http://vphill.com/journal/post/5569/
  • http://vphill.com/journal/post/5908
  • 1,162,491 unique headings
  • 1,827,276 records in the DPLA that have no subjects or keywords
“Metadata Analytics, Visualization, and Optimization: Experiments in statistical analysis of the Digital Public Library of America (DPLA) “

• Research by Corey Harper, NYU
• To be published in upcoming Code4Lib Journal
• Visualization of metadata
• Natural language processing
• Google analytics & usage data
• Predictive analytics
Collection Use and SEO

“Getting Found: Search Engine Optimization for Digital Repositories”

• IMLS funded work: Montana State, OCLC, and University of New Mexico

• SEO Cookbook – http://www.clir.org/pubs/reports/pub165

• DLF Assessment Group - https://www.diglib.org/groups/assessment/

• How can DPLA help with collection Assessment?
DPLA: Dataset of Value

• Research Community and iSchools
• Shared Research Agenda

  – Metadata Models

  – Information Architecture / Knowledge Organization

  – Assessment / Usage / Usability

  – Collections Data
DPLA Collections Data

• Understanding the “Aboutness”

• Where are voices silent in our national narrative?

• How do we correct the bias?

• How do we build workflows that build equitable collections?
Cultural Heritage for Social Change

- *Documenting the Now*
  - [http://mith.umd.edu/introducing-documenting-the-now/](http://mith.umd.edu/introducing-documenting-the-now/)
  - [http://www.docnow.io/](http://www.docnow.io/)

- Community Archives

- “Our Story” Knight News Challenge Grant
“For archivists to abstain from cultural awareness and criticism is tantamount to professional irresponsibility.”
~ Brien Brothman

“The library serves democracy by providing a neutral forum in which all types of information and all points of view find equal voice.”
~ Patricia Wilson Berger
“Let us redouble our efforts to ensure that our repositories become places of inclusion that celebrate the diversity of our society and the historical record. Let us strive to promote free and equitable access to the primary historical record that promotes understanding of the truth and that fights against ignorance and misrepresentation of the American experience...

When we preserve the records and share the histories of our marginalized communities, we are working toward alleviating the fear and hate that led to the Orlando shooting and to so many other violent acts that our country has witnessed.”

~ Dennis Meissner

https://offtherecord.archivists.org/2016/06/15/statement-on-the-orlando-shooting/
“I see the building of a new library as one of the greatest examples of what humans can do together to extend the light against the darkness.”

~ Dan Cohen
DPLA: Going Forward
Completing the Network
The Digital Public Library of America (DPLA), Stanford University and DuraSpace are partnering to extend the existing Hydra project codebase and its vibrant and growing community to build, bundle, and promote a feature-rich, robust, flexible digital repository that is easy to install, configure, and maintain.

This next-generation repository solution – “Hydra-in-a-Box” – will work for institutions large and small, incorporating the capabilities and affordances to support networked resources and services in a shared, sustainable, national platform. The overall intent is to develop a digital collections platform that is not just “on the web,” but “of the web.”

With funding from the Institute of Museum and Library Services, the 30-month collaborative project kicked off in May 2015; product design and software development are now underway.

http://hydrainabox.projecthydra.org
Site-wide Support of HTTPS

Library Digital Privacy Pledge –
https://libraryfreedomproject.org/ourwork/digitalprivacypledge/

https://dp.la/info/2016/06/01/dpla-launches-full-support-for-https/
Helping all children discover a love of reading

Find out more
Find out More


• FaceBook: https://www.facebook.com/digitalpubliclibraryofamerica/

• Twitter: @dpla
DPLA: More than a Portal

- a voice for best practices in GLAMs
- a means for advocacy and implementation
- demonstrate the benefit of openness and reuse
- embody professional values
“The Mission of *Librarians* is to Improve Society through Facilitating Knowledge Creation in their Communities”

R. David Lankes
*Atlas of New Librarianship.*

[http://www.newlibrarianship.org/wordpress/]
Thank you!
Questions?
Comments?
rachel@dp.la
Twitter: @rlfrick