Technology and Metadata Project Showcase

Moderators:
Katie McCormick and Bradley Daigle
Visualizing Funding for Libraries

Sarah Rice
Forefront
VISUALIZING FUNDING FOR LIBRARIES

Sarah Rice, Forefront Librarian
Track and Find FUNDING FOR LIBRARIES
A data tool and capacity building training to help libraries and their supporters thrive

EXPLORE DATA TOOL

EXPLORE TRAINING
Visualizing Funding For Libraries: A Project of Foundation Center with support from the Knight Foundation
### Subject Area

- **Arts and culture**
  - Cultural awareness
  - Folk arts
  - Historical activities
  - Humanities
  - Performing arts
  - Public arts

- **Community and economic development**

- **Education**

- **Environment**

- **Health**

- **Human rights**

- **Museums**
  - Children's museums
  - Ethnic museums
  - History museums
  - Maritime museums
  - Natural history museums
  - Science museums
  - Sport and hobby museums

### Filters

- **Population Served**
  - Select Population Served

- **Support Strategy**
  - Select Support Strategy

- **Organization Name**
  - Enter Organization Name

- **Grant Amount Range**
  - From USD
  - To USD

- **Grant Years**
  - 2006
  - 2007
  - 2008
  - 2009
  - 2010
  - 2011
  - 2012
  - 2013
  - 2014
  - 2015
  - 2016
  - 2017
  - 2018
  - 2019
  - 2020

- **Funder Type**
  - Select Funder Type

- **Filter by Keyword**
  - Enter keyword(s)
Funders funding organizations based in the United States in support of Digital collections starting in year(s) 2006 to present:

<table>
<thead>
<tr>
<th>Funders</th>
<th>City</th>
<th>State</th>
<th>Country</th>
<th>Dollar Value of Grants</th>
<th>Number of Grants</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Andrew W. Mellon Foundation</td>
<td>New York City</td>
<td>NY</td>
<td>United States</td>
<td>$25.3 Million</td>
<td>16</td>
</tr>
<tr>
<td>Carnegie Corporation of New York</td>
<td>New York City</td>
<td>NY</td>
<td>United States</td>
<td>$6 Million</td>
<td>3</td>
</tr>
<tr>
<td>Arcadia Fund</td>
<td>London</td>
<td>England</td>
<td></td>
<td>$2.5 Million</td>
<td>1</td>
</tr>
<tr>
<td>Alfred P. Sloan Foundation</td>
<td>New York City</td>
<td>NY</td>
<td>United States</td>
<td>$1.9 Million</td>
<td>1</td>
</tr>
<tr>
<td>Recipient Name</td>
<td>City</td>
<td>State</td>
<td>Country</td>
<td>Dollar Value of Grants</td>
<td>Number of Grants</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>-------</td>
<td>-------------</td>
<td>------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Library of Congress</td>
<td>Washington DC</td>
<td>DC</td>
<td>United States</td>
<td>$5.5 Million</td>
<td>5</td>
</tr>
<tr>
<td>Letterform Archive</td>
<td>San Francisco</td>
<td>CA</td>
<td>United States</td>
<td>$17,514</td>
<td>4</td>
</tr>
<tr>
<td>Council on Library and Information Resources</td>
<td>Washington DC</td>
<td>DC</td>
<td>United States</td>
<td>$21.6 Million</td>
<td>3</td>
</tr>
<tr>
<td>Digital Public Library of America</td>
<td>Boston</td>
<td>MA</td>
<td>United States</td>
<td>$4.4 Million</td>
<td>2</td>
</tr>
<tr>
<td>Ny Public Library</td>
<td>New York City</td>
<td>NY</td>
<td>United States</td>
<td>$2 Million</td>
<td>2</td>
</tr>
<tr>
<td>Abilene Library Consortium</td>
<td>Abilene</td>
<td>TX</td>
<td>United States</td>
<td>$1.4 Million</td>
<td>2</td>
</tr>
<tr>
<td>John F. Kennedy Library Foundation</td>
<td>Boston</td>
<td>MA</td>
<td>United States</td>
<td>$250,000</td>
<td>2</td>
</tr>
<tr>
<td>Newberry Library</td>
<td>Chicago</td>
<td>IL</td>
<td>United States</td>
<td>$58,000</td>
<td>2</td>
</tr>
</tbody>
</table>
# Grant Details

**Funder**
Alfred P. Sloan Foundation  
New York City, NY United States

**Recipient**
Digital Public Library of America  
Boston, MA United States

## Basic Grant Information
- **Total Amount:** $1,901,709  
- **Year Awarded:** 2015  
- **Grant Duration:** 2 years and 6 months  
- **Program Area:** Digital Information Technology

## More Details
- **Primary Subject:** Digital collections  
- **Additional Subjects:** Libraries  
- **Population Served:** Not Specified  
- **Support Strategy:** Network building and collaboration. Pilot programs; Online media

## Grant Description
To complete Nationwide Service Hub Network and to pilot an e-books distribution program
Funding pathways to organizations based in United States; starting in year(s) 2006 to present starting with Digital Public Library of America

Digital Public Library of America...
Bill & Melinda Gates Foundation
Alfred P. Sloan Foundation
The Elkes Foundation
University of Washington...
Cornell University
NATIONAL YIDDISH BOOK CENTER
University of California at...
Grants where funder is The Paul G. Allen Family Foundation, to organizations based in United States; starting in year(s) 2006 to present

<table>
<thead>
<tr>
<th>Funder Name</th>
<th>Recipient Name</th>
<th>Year</th>
<th>$ Dollar Value of Grant</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>University of Washington Foundation</td>
<td>2012</td>
<td>$3 Million</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Library Foundation</td>
<td>2007</td>
<td>$400,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Seattle University</td>
<td>2008</td>
<td>$250,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Kitsap Regional Library Foundation</td>
<td>2012</td>
<td>$200,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Washington State Library</td>
<td>2012</td>
<td>$200,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Library Foundation</td>
<td>2013</td>
<td>$196,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Library Foundation</td>
<td>2012</td>
<td>$190,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>PIERCE COUNTY LIBRARY FOUNDATION</td>
<td>2013</td>
<td>$175,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>Tacoma Public Library</td>
<td>2010</td>
<td>$150,000</td>
<td></td>
</tr>
<tr>
<td>The Paul G. Allen Family Foundation</td>
<td>PIERCE COUNTY LIBRARY FOUNDATION</td>
<td>2012</td>
<td>$150,000</td>
<td></td>
</tr>
</tbody>
</table>
Organization Name: Enter Organization Name
Grants to organizations based in United States; in support of Digital collections; starting in year(s) 2006 to present; 115 of 115 links
Resources

Grantseeking for Libraries
(Self-paced eLearning)
https://bit.ly/2Qxy2v8

Grantseeking for Digital Libraries
(Member Webinar recording)
https://pro.dp.la/events/workshops

Visualizing Funding for Libraries
(Data tool)
http://libraries.foundationcenter.org/
Thank you

Questions?

srice@myforefront.org
https://guides.myforefront.org
Metadata Cleanup in the Pacific Northwest

Evan Robb
Washington State Library
Metadata Cleanup in the Pacific Northwest
Building Library Capacity with Grant Support

Evan Robb
Digital Projects Librarian
Washington State Library
evan.robb@sos.wa.gov
2016 – 2018 LSTA Grants

Washington State Library

- $5,000 grants to libraries.
- Metadata assessment, recommendations, training & support.

Orbis Cascade Alliance

- Harvester development.
- Formal training.
- Project Metadata Consultant.
- Documentation and tools.

Evaluate the Collection's Metadata

- Metadata Readiness Checklist
- Metadata Cleanup Decision Matrix
- Assess whether your metadata meets requirements set by the Dublin Core Best Practices, plan where to clean up your metadata (in your DAM/IIR, with OpenRefine, or in the Harvester).

- Required Data checker and Dublin Core Mapping Checker
- Check whether your records are missing required metadata.
- See what kind of metadata is mapped to each Dublin Core field, catch mapping errors, and identify metadata for cleanup.

- Metadata Enrichment Checklist
- Thinking of enriching in addition to cleanup? This checklist will help you decide what, and how much, enrichment is suitable.

- Alliance Metadata Standards
- Dublin Core Best Practices
- Forming Creators & Contributors According to ECA

Determine and Apply Rights Statements

- Rights Statements in Digital Collections
- Rights Statements Decision Matrix

Remediate your Metadata

- Batch Editing CONTENTdm Records
- Batch Editing Digital Commons Records
- Batch Editing Dspace Records
- Batch Editing Omeka Records

- OpenRefine for Metadata Cleanup
- Do batch cleanup on records exported from your local system

- OpenRefine for Metadata Enrichment
- Match metadata values to controlled vocabulary terms

https://www.orbiscascade.org/digital-collections-documentation/
Partner Institutions

Grant-funded cleanup (WA)

• Evergreen State College
• Port Angeles Public Library
• Seattle Public Library
• Western Washington University
• Whitman County Library
• Whatcom County Library System

Training & consulting participants

• Densho
• Everett Public Library
• Multnomah County Library
• Oregon Historical Society
• Spokane Public Library
• Tacoma Public Library
• ~26 Alliance libraries

Northwest Postcard Collection Metadata
Tacoma Public Library
April 2017

Recommendations
Add the following fields to the Northwest Postcard Collection metadata schema:
- Contributor
- Date (human readable)
- Format (digital)
- Language
- Rights (free-text)
- Rights (standardized)
- Spatial Coverage
- Transcription
- Type, after renaming current Type field to Genre

Consolidate all printed text into the new field “Transcription” and remove the following fields:
- Back caption
- Front caption
- Distributor
- Printer
- Publisher
- Notes.

Review of Existing Metadata

Title
- All records have one title value.
- Titles are structured in a consistent format.
  - Name of the subcollection, image number within the subcollection
- Recommendations: Devise a brief title that includes an indication of its subject (e.g., names of persons, corporate bodies, objects, activities, events, geographical area and date). It may include the front or back caption, some examples:
  - GRIPP-03 would become: Olympic Min. Union City, ca. 1907
  - GRIPP-02 would become: Scenic View of Hood Canal Washington, ca. 1910
Trainings

- Data and Quality at DPLA.
- Best Practices for Dublin Core Metadata.
- Metadata Readiness.
- Metadata Cleanup.
- Metadata Enrichment.
- Copyright and Risk Management.
- Assigning Standardized Rights Statements (RightsStatements.org)
- Creating Free-text Rights Statements.
- RDA for Personal Names, Devised Titles.
Remediated!

Cleanup Grant Recipients
- ~20,500 remediated records ($28K).
- Avg. ~4 minutes per record for diverse collections.
- Increased confidence with batch and automated tools/processes.

Orbis Cascade Alliance
- ~67,000 remediated records in first year (currently ~100,000 (?)).

Ready for harvest.
Thank you!

Evan Robb
Digital Projects Librarian
Washington State Library
evan.robb@sos.wa.gov
Bridge2Hyku: Progress on a Migration Toolkit

Todd Crocken
University of Houston Libraries
BRIDGE2HYKU: PROGRESS ON A MIGRATION TOOLKIT
GOALS > TIMELINE > PROGRESS > CONTRIBUTE
TODD CROCKEN - DPLAfest - APRIL 2019
BASICS

- 2 Year IMLS grant project (lg-70-17-0217-17)
- Based at University of Houston with partner institutions: University of Victoria, University of Miami, Texas Digital Library, and IUPUI
- Advisors/Stakeholders: Stanford, Indiana, DPLA, Duraspace.
GOALS

● Build a website with content surrounding migration strategies and best practices
● Create open source software for migration
  ○ Exporting out of CONTENTdm
  ○ Importing into Hyrax/Hyku
● Community support and sustainability
TIMELINE

Phase 1
Oct 2017 - Apr 2018
Conduct need assessment (metadata and system requirements)

Phase 2
May 2018 - Apr 2019
Develop and document B2H Toolkit and website

Phase 3
May 2019 - Sept 2019
PHASE ONE DONE

- Digital Collections Survey done
  - What does a “typical” institutional digital library look like?
- Website launched
THE WEBSITE

- Migration best practices
- Online resources
  - Blogposts about specific issues
  - Links to other useful sites
  - Migration framework
- Toolkit instructions and info
- End of Phase 1&2 Reports
- Hyku features
PHASE TWO WRAPPED UP

● **Software tools developed through 15+ dev sprints**
  ○ Evaluated existing migration tools
    ■ University of Victoria’s `cdm_migrator` gem
  ○ Generated MVP use cases/user stories for migration toolkit
    ■ 6 week working group steered by Anne Washington
  ○ Iterative development of CDM Bridge and HyBridge - Summer 2018
    ○ Beta/Release Candidates of both tools available now!

● **Toolkit Pilot Test completed**
  ○ 2 Months of testing migration paths through the B2H tools
  ○ Report now available

● **Blogposts and build-out of migration content on website**
  ○ `bridge2hyku.github.io/blog`

- Desktop Electron-app
- Works with hosted CONTENTdm
- Metadata crosswalk functionality
- **Default Compatibility with DPLA’s MAP v4**
- Export csv and files for various uses
<table>
<thead>
<tr>
<th>Export Field</th>
<th>Item Level</th>
<th>CDM Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Required</td>
<td>Title</td>
</tr>
<tr>
<td>Creator</td>
<td>Required</td>
<td>Creator (LCNAF)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creator (HOT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creator (ULAN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creator (Local)</td>
</tr>
<tr>
<td>Keyword</td>
<td>Required</td>
<td>Subject:Topical (Local)</td>
</tr>
<tr>
<td>Rights statement</td>
<td>Required</td>
<td>Use and Reproduction</td>
</tr>
<tr>
<td>Contributor</td>
<td></td>
<td>Contributor (LCNAF)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Contributor (HOT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Contributor (ULAN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Contributor (Local)</td>
</tr>
<tr>
<td>Abstract or Summary</td>
<td></td>
<td>Description</td>
</tr>
<tr>
<td>License</td>
<td></td>
<td>-- Select a field --</td>
</tr>
<tr>
<td>Publisher</td>
<td></td>
<td>-- Select a field --</td>
</tr>
</tbody>
</table>

- Ruby Gem for Hyrax and Hyku
- Creates easy to use tab for bulk ingest
- Allows collection-specific ingest
- **Seamlessly integrated with CDM Bridge’s output**
Phase 3?

- Workshops on Digital Repository Selection and Migration
  - Digital Initiatives Symposium - April 29-30th, San Diego
  - Joint Conference on Digital Libraries - June 2-6th, Urbana-Champaign

- Open beta testing of software
  - Available to download and use now!
  - Let us know about any bugs you encounter or features you’d like, we have dev hours devoted to any fixes/updates the community brings up
CONTRIBUTE

- Join CONTENTdm Migration IG on Samvera’s slack: #hydra-migrations
- Have a migration story for a blogpost or want to help with testing/documentation? Contact Todd Crocken (tcrocken@uh.edu)
- Follow us on Twitter! @BRIDGE2Hyku
- Check out our github: https://github.com/Bridge2Hyku
THANKS!

Background from unsplash.com

Abstract: The Bridge2Hyku Project, an IMLS National Leadership/Project Grant (LG-70-17-0217-17), has finished its second phase and has produced a multifaceted migration toolkit. We will provide information about our toolkit, including software (CDM Bridge, HyBridge) and our migration guide, and how it will help those considering moves from legacy systems to open source platforms compatible with the DPLA.
Aggregating with Eliza and Barb: creating user personas for software development at PA Digital

Leanne Finnigan
Temple University, PA Digital
Aggregating with Eliza and Barb

Creating user personas for software development at PA Digital

Leanne Finnigan
metadata coordinator
PA Digital
Samvera-based aggregator

**DPLA Aggregator**

**Editing OAI seed**

Note: If the OAI implementation for this seed uses sets, enter one set at a time (one set per OAI seed).

- **Name**
  - 2nd Geological Survey (PA)
  - Note: not for end user display

- **Endpoint url**
  - http://psa.powertoolibrary.org/oai/oai.php

- **New endpoint url**

- **Metadata prefix**
  - oai_qdr

- **Set**
  - gsabc

- **Collection name**
  - 2nd Geological Survey (PA)
  - Note: For end user display
Samvera-based aggregator
Eliza coordinates metadata reviews. She has an expert level knowledge of metadata and OAI. She prefers a lot of control and the ability to work autonomously without developer involvement in the day-to-day.

Jo is on the development team. They are assigned tickets to make fixes but often have to grapple with a possible ripple effect that the changes may have on the rest of the code.

Greta works at DPLA and coordinates all of the hub’s quarterly harvests. Occasionally, she will point out errors that cannot be fixed on the DPLA side. She often asks for input and collaborates with hubs to make the metadata ingestion process painless.
funnel cake image from nathaliewilson via flickr
Many thanks to:

Rachel Appel, Temple University
Linda Ballinger, Pennsylvania State University
Doreva Belfiore, HSLC
Michael Carroll, Temple University
Jasmine Clark, Temple University
Pam Dinan, IDS-PA
Will Echevarria, Free Library of Philadelphia
Ana Enriquez, Pennsylvania State University
Bill Fee, State Library of Pennsylvania
Leanne Finnigan, Temple University
Gabe Galson, Temple University
Brandy Karl, Pennsylvania State University
Alice Lubrecht, State Library of Pennsylvania
Joe Lucia, Temple University
David McKnight, University of Pennsylvania
Glenn Miller, State Library of Pennsylvania
Chad Nelson, Temple University
Steven Ng, Temple University
Maryam Phillips, HSLC
Stefanie Ramsay, Temple University
Katy Rawdon, Temple University
Martina Soden, Scranton Public Library
Matthew Strauss, Senator John Heinz History Center
Holly Tomren, Temple University
Cathy Wilt, PALCI
Jennifer Wright, Free Library of Philadelphia
And all our other contributors...

This project is made possible by a grant from the Institute of Museum and Library Services as administered by the Pennsylvania Department of Education through the Office of Commonwealth Libraries, and the Commonwealth of Pennsylvania, Tom Wolf, Governor.
Introducing Pyrepxox: A Repox Client for Python

Mark Baggett

University of Tennessee
Introducing Pyrepox

A Repox Client for Python

Mark Baggett
Head, Digital Initiatives
University of Tennessee
Repox and Pyrepx in a Nutshell

- **Repox**: A software tool for managing metadata aggregation
  - Supports numerous protocols (OAI-PMH, File System Uploads, Z39.50, etc.)
  - Old and not supported
  - People Still Use it

- **Pyrepx**: An implementation of Repox’s REST API in Python
  - Makes automating things in Repox easy
  - Extends some of the built-in APIs in Repox
Requirements

- Repox 3.0.0 and up ([https://github.com/europeana/repox](https://github.com/europeana/repox))
- Currently Python 3.7 and up (refactoring for 3.5 and up)
- Update Repox's Swagger UI class (instructions here: [https://github.com/europeana/REPOX/wiki/Using-Swagger-UI](https://github.com/europeana/REPOX/wiki/Using-Swagger-UI))
Getting Started

```bash
pip install repox

```
List Aggregators and Providers

```python
>>> r.list_all_aggregators()
['TNDPLAr0']

>>> r.get_list_of_providers('TNDPLAr0')
['CountryMusicHallofFamer0', 'CrossroadstoFreedomr0', 'KnoxPLr0', 'memphispublicr0', 'MTSUr0', 'nashvillepublicr0', 'tslar0', 'UTKr0', 'utcr0']
```
Get and Update Metadata about a Provider

```python
>>> r.get_provider('utcr0')
{'id': 'utcr0', 'name': 'University of Tennessee at Chattanooga',
'countryCode': 'usa', 'description': 'contact: Carolyn Runyon
platform: ContentDM', 'nameCode': 'utc', 'homepage':
'http://cdm16877.contentdm.oclc.org/oai/oai.php', 'providerType':
'LIBRARY', 'email': 'carolyn-runyon@utc.edu'}

>>> r.update_provider("utcr0",
    homepage="http://dloai.lib.utk.edu/cgi-bin/XMLFile/dlmodsoai/oai.p1",
    email="mbagget1@utk.edu")
200
```
Get List of Sets or Details about a Set

```python
>>> r.get_list_of_sets_from_provider("utcr0", verbose=False)
```
Get Record Counts

```python
>>> r.count_records_in_dataset('nr')
7518

>>> r.count_records_from_provider('utcr0')
15919
```
Or Get Statistics About Your Entire Instance

```python
>>> r.get_statistics()
{
    "repox-statistics": {
        "@generationDate": "2019-04-08 15:32:52 EDT",
        "dataSourcesIdExtracted": "0",
        "dataSourcesIdGenerated": "19",
        "dataSourcesIdProvided": "193",
        "aggregators": "1",
        "dataProviders": "9",
        "dataSourcesOai": "193",
        "dataSourcesZ3950": "0",
        "dataSourcesDirectoryImporter": "19",
        "dataSourcesMetadataFormats": [
            {"metadataFormat": "mods", "dataSources": "52", "records": "30366"},
            {"metadataFormat": "oai_dc", "dataSources": "86", "records": "160203"},
            {"metadataFormat": "oai_qdc", "dataSources": "58", "records": "31070"},
            {"metadataFormat": "xoai", "dataSources": "16", "records": "4366"}
        ],
        "recordsAvgDataSource": "1066.0613",
        "recordsAvgDataProvider": "25111.666",
        "countriesRecords": {
            "countryRecords": [
                {"@country": "usa", "records": "105854"},
                {"@country": "de", "records": "120151"}
            ]
        },
        "recordsTotal": "226005"
    }
}
Schedule Harvests

```python
>>> r.schedule_weekly_harvest('p16877coll1', 'Sunday', time='04:00')
201
```

```python
>>> r.schedule_harvest('bcpl', frequency='XMONTHLY', xmonths=2, time='01:00', date='01/12/2019')
201
```
Get Recently Harvested Sets

```python
>>> r.get_recently_ingested_sets_by_aggregator('TNDPLAr0', '2019-03-10')
Create and Delete Things

```python
>>> r.create_provider("TNDPLAr0", {"id": "utc", "name": "UT Chattanooga", "country": "United States", "countryCode": ","description": "OAI Sets from the University of Tennessee, Chattanooga", "nameCode": "utc", "homepage": "http://cdm16877.contentdm.oclc.org", "providerType": "LIBRARY", "email": "carolyn-runyon@utc.edu"})
200

>>> r.delete_automatic_harvesting_task('nr', 'nr_3')
200

>>> r.delete_dataset('nashville_test')
200

>>> r.delete_provider('utcr0')
200
```
Some Practical Applications

University of Tennessee at Chattanooga

Details

Email: carolyn-runyon@utc.edu
Description: contact: Carolyn Runyon platform: ContentDM
OAI endpoint:
Total records in Repox: 15919
Page last updated: 2019-03-19

Datasets

p16877coll1

- Record Total: 91
- In DPLA?: True
- Metadata format: oai_qdc
- Last ingest date: 03/13/2019 10:20:51
- Dataset Type: OAI
- OAI Set: p16877coll1
- Records Being Sent to DPLA: 71
Problems / Work to Do

1. Upload XSLT to the filesystem
2. Associate stylesheets with set
3. Delete records?
4. Master list of Todos here:
Learn More

- GitHub (https://github.com/markpbaggett/pyrepox)
- ReadTheDocs (https://pyrepox.readthedocs.io)
- PyPi (https://pypi.org/project/repox/)
Process Assessment as a Cultural Shift

Clara Hudson
Virginia Beach Public Library
Process Assessment as a Cultural Shift

Clara Hudson
Virginia Beach Public Library
Where It Started....

- Two units
  - Cataloging
  - Collection management
- Working in opposition
- Real Estate
- Three month backlog
- Budget cuts
  - Materials
  - Staff
The Work

• Taskforce to define the work
• Redundancy, gaps, bottlenecks
• Budgeting for efficiency
• What’s in a name
• Moving in together
• Job descriptions
• Backlog
• Documentation
• Training
  • More training
Facilitated Process for Improvement

Task: Improve the efficiency & ROI for acquiring & processing Library Department’s materials collection.

Current State:
- We work separately
- There is baggage (history)
- There are individual performance issues
- Perfectionism & Professionalism - both of these "bests" get in the way of "better."
- There is resistance to change
- We have different priorities than ComIT
- Lack of empathy (understanding of the work)
- Lack of validation
- Processes are not streamlined
- Processes are not collaborative
- No time to analyze data we collect
- Perception - customers aren’t getting materials as quickly as they want them (immediately)
- Changes in publishing world
- Changes in technology
- Ever changing formats
- Cataloging rules changing

Key Stakeholders:
- GS/CM Staff
- Automated Services
- Library Staff
- Mgt Staff of GS/CM
- ComIT
- Vendors
- Service Development Manager
- Citizens
- Customers

Action Steps:
- Frame the Issue
- Create opportunities for cross-training
- Reduce Current Bottlenecks
- Staff Awareness
- Form and launch a team
- Analyze Workflow
- Research and Explore alternatives & opportunities

Measures for success:
- Better Customer Service
- Improve Efficiency, Workflow & Productivity
- Broaden Skill Sets
- Increased Collaboration Across Units

Future State:
- Shared ownership of collections by all staff
- Materials are acquired and processed in a timely fashion
- ROI improved
- Support Services aligned with Customer Needs
- Staff Relationships improve

Bottom-line, what we are doing is: Improving the Efficiency & Return on Investment in Collection Management and Cataloging Services. VBPL Initiative 9A: Maximize the Library’s effectiveness through efficient use of resources and responsible stewardship.
Data

- Defining
- Sourcing
- Analysis
- Utilization
- Evidence based acquisition
- Actual demand
- Threshold of work
**Project Management**

- Opening day collection
- Floating collections
  - Planning and collaboration
  - Collection clean-up
  - The story of where collections want to live
- Assess, Assess and Assess again
- Process assessment x3
- Balance and redistribution
- Inventory and Weeding
- Taking Ownership of Collection Management

*We did more with less! We did it better than before.*
Culture

- Teamwork
- Team building
- Change management
- Public service staff
  - Customers
• Continued assessment
• Process improvement
• Job reclassification
• Who does what, when, how and WHY
• Training, cross training, collaboration, redundancy for back-up

We are still process improving, and have been able to voluntarily transfer vacant positions where they are needed based on changing use patterns, collection formats, and leveraging technology…based on evidence.
Clara Hudson
Support Services Administrator
Virginia Beach Public Library
chudson@vbgov.com
757-385-0171
The Illinois Digital Heritage Hub Type Metadata Project: Increasing User Access and Record Visibility

Joshua Lynch

University of Illinois, Urbana-Champaign
The Illinois Digital Heritage Hub Type Metadata Project

Increasing User Access and Record Visibility

Joshua Lynch and Jessica Gibson
Better Metadata Helps Contributors, DPLA, and users

Contributor provides consistent type metadata for the DPLA’s catalog

User finds resources and provides traffic to local contributors

DPLA provides facets for DCMI Type metadata, linking to other collections and increasing interactivity

DPLA User
Type Metadata: Lowest Completion Rates of any Highly Recommended Field
Collecting Data

Python Script to gather JSON from DPLA’s API:

- Item ID
- Collection
- Provider
- Original Record Type
- Source Resource Type
Analyzing Data with OpenRefine

Analyzed JSON in OpenRefine:

- Faceting
- Facet counts
- Blankdown
- Create new column based on sum of CSVs
## Analysis Findings: More than 550 type values

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>505</td>
<td>Text; 2nd Edition, proof signature</td>
</tr>
<tr>
<td>506</td>
<td>Text; advertisement page</td>
</tr>
<tr>
<td>507</td>
<td>Text; book; Images</td>
</tr>
<tr>
<td>508</td>
<td>Text; colored illustrations</td>
</tr>
<tr>
<td>509</td>
<td>Text; Envelopes</td>
</tr>
<tr>
<td>510</td>
<td>Text; final proof</td>
</tr>
<tr>
<td>511</td>
<td>Text; Image</td>
</tr>
<tr>
<td>512</td>
<td>Text; Image;</td>
</tr>
<tr>
<td>513</td>
<td>Text; Image; Sound;</td>
</tr>
<tr>
<td>514</td>
<td>Text; Image; Text, Image</td>
</tr>
<tr>
<td>515</td>
<td>Text; Ledger Page</td>
</tr>
<tr>
<td>516</td>
<td>Text; octavo</td>
</tr>
<tr>
<td>517</td>
<td>Text; offprint</td>
</tr>
</tbody>
</table>
How We Fixed It

IDHH

Communication

Contributors

Communication

DPLA

DIGITAL PUBLIC LIBRARY OF AMERICA
How We Fixed It

- Inform the DPLA
- Transform with XSLT
- Type metadata guide: https://bit.ly/2TPsCkt

Contributors
How We Fixed It

- Use controlled vocab
- Avoid non-standard usage

Contributors
How We Fixed It

IDHH

- Documentation and communication
- Ingestion better prepared for data

Contributors

DIGITAL PUBLIC LIBRARY OF AMERICA
Results

Before

Refine your search

<table>
<thead>
<tr>
<th>Type</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>image</td>
<td>58,772</td>
</tr>
<tr>
<td>text</td>
<td>1,547</td>
</tr>
<tr>
<td>sound</td>
<td>1,062</td>
</tr>
<tr>
<td>moving image</td>
<td>20</td>
</tr>
<tr>
<td>physical object</td>
<td>1</td>
</tr>
</tbody>
</table>

After

Refine your search

<table>
<thead>
<tr>
<th>Type</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>image</td>
<td>226,918</td>
</tr>
<tr>
<td>text</td>
<td>65,600</td>
</tr>
<tr>
<td>sound</td>
<td>2,820</td>
</tr>
<tr>
<td>physical object</td>
<td>2,148</td>
</tr>
<tr>
<td>moving image</td>
<td>633</td>
</tr>
</tbody>
</table>
Results

Before

<table>
<thead>
<tr>
<th>Metadata Completeness</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Description</td>
<td>83%</td>
<td>83%</td>
</tr>
<tr>
<td>Creator</td>
<td>56%</td>
<td>56%</td>
</tr>
<tr>
<td>Type</td>
<td>20%</td>
<td>93%</td>
</tr>
<tr>
<td>Language</td>
<td>45%</td>
<td>45%</td>
</tr>
<tr>
<td>Spatial</td>
<td>63%</td>
<td>64%</td>
</tr>
<tr>
<td>Subject</td>
<td>82%</td>
<td>82%</td>
</tr>
<tr>
<td>Date</td>
<td>87%</td>
<td>86%</td>
</tr>
<tr>
<td>Rights</td>
<td>81%</td>
<td>81%</td>
</tr>
<tr>
<td>Standardized Rights</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td>Preview</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

After
Type Metadata Paper Link

https://bit.ly/2WcBixV
LINEAGE: An artificially intelligent visual discovery engine

Noya Kohavi
Brown Institute & Columbia University Graduate School
Lineage

A fully visual discovery engine and collection explorer

Noya Kohavi
April 2019
Existing services

**Google reverse image search:**
Helps identify the object, finds more of the same

**Museum collection databases:**
Textual search means you have to already know what you’re looking for

---

Noya Kohavi

April 2019
Original Usage: contextualizing fashion

Input

NY Fashion Week
Street fashion, Fall 2018

Lineage output

Yves Saint Laurent
Ensemble, S/S 1990
Met Costume Institute

Ellsworth Kelly
White and Black, 1991
MoMA

Ellsworth Kelly
Light Reflection on Water, 1950
MoMA

Noya Kohavi
April 2019
LINEAGE  A fully visual discovery engine and collection explorer

CÉLINE
Fall 2018 RTW

Nina Ricci
Pre-fall 2018

Issey Miyake
ensemble, ca 1993
Met Costume Institute

Whitney Museum of American Art

Clifford Still
Untitled, 1956
Whitney Museum of American Art

Sonia Rykiel
Spring 2019

Issey Miyake
Dress, S/S 1994
Met Costume Institute

Lineage
Automatic art + design context | Feb. 2019

Noya Kohavi
Automatically generated similarity matches created by Lineage:

Georges Vantongerloo
*Relation of Lines and Colors*
Paris 1939

Cy Twombly
*The Italians*
Rome, January 1961

Raymond Parker
*Untitled*
1979
Automatically generated similarity matches created by Lineage:

Moses Soyer
*Artist and Model*
1962

Lucian Freud
*Portrait of a Woman*
1949
Automatically generated similarity matches created by Lineage:

Martin Baer
*Parrot Tulips*
1939

Jake Berthot
*Room*
1979
Automatically generated similarity matches created by Lineage:

Sean Scully
*Big Grey Robe*
2002

Kurt Roesch
*Bones on the Table*
1939
Automatically generated similarity matches created by Lineage:

Pablo Picasso
*Woman Dressing Her Hair*
Royan, June 1940

Bedri Rahmi
*The Chain*
1962
Automatically generated similarity matches created by Lineage:

Philip Pearlstein
Two Female Models in the Studio
1967

Kenzo Okada
Number 3
1953
Automatically generated similarity matches created by Lineage:

John D. Graham
Self-Portrait
1943

André Derain
Charing Cross Bridge
London 1905-06
Automatically generated similarity matches created by Lineage:

Marcel Duchamp
*Landscape*
Neuilly, January-February 1911

Hans Hofmann
*Spring*
1944-45 (dated on reverse 1940)
Automatically generated similarity matches created by Lineage:

- Joan Miró, *The Hunter (Catalan Landscape)*, Montroig, July 1923-winter 1924
- Max Oppenheimer (MOPP), *The World War*, 1916
Automatically generated similarity matches created by Lineage:

Umberto Boccioni
*The Riot*
1911 or after

Stanton Macdonald-Wright
*Synchromy*
1917
Automatically generated similarity matches created by Lineage:

Amedeo Modigliani
Anna Zborowska
1917

Édouard Vuillard
Dinnertime
c. 1889
Automatically generated similarity matches created by Lineage:

Francesco Clemente
Perseverance
1981

Pat Steir
Nabaye Tabastan for L.T. & R.
1974
Automatically generated similarity matches created by Lineage:

Rubem Valentim
2x-2
1960

Kerry James Marshall
*Untitled (policeman)*
2015
Automatically generated similarity matches created by Lineage:

Joseph Stella
*First Light*
c. 1928

Gino Severini
*Armored Train in Action*
1915
Automatically generated similarity matches created by Lineage:

Bryan Wynter
*Meeting Place*
1957

Sue Coe
*Woman Walks into Bar - Is Raped by Four Men on the Pool Table - While 20 Watch*
1983
Automated similarity matches created by Lineage:

Jo Baer
*Double-Bar, Grey (Green Line)*
1968

Alfredo Hlito
*Cromatic Rhythms III*
1949
Automated similarity matches created by Lineage:

Laura Owens
*Untitled*
2013

A. E. Gallatin
*Forms and Red*
December 1949
Automated similarity matches created by Lineage:

Abraham Palatnik  
*Progression 91-A*  
1965

Ben Vautier  
*Art*  
1974
Automatically generated similarity matches created by Lineage:

Robert Barry
*Untitled*
August 1967

Daniel Buren
*One Painting in Four Elements for One Wall*
1969-77
Automatically generated similarity matches created by Lineage:

**George Maciunas**  
*Untitled*  
Early 1950's

**Gerhard Richter**  
*Cityscape*  
1970
Automatically generated similarity matches created by Lineage:

Odilon Redon
*Trees in Bièvres*
after 1900

Odilon Redon
*Underwater Vision*
c. 1910
Automatically generated similarity matches created by Lineage:

Jasper Johns
*Diver*
1962-63

Ben Vautier
*Flux Painting*
c. 1966

David Webster
*Cerebellar Islands*
1999
Automatically generated similarity matches created by Lineage:

**Richard Hamilton**  
*Glorious Techniculture*  
1961-64

**Jess**  
*Ex. 4-Trinity's Trine*  
1964
Automatically generated similarity matches created by Lineage:

Willem de Kooning
*Untitled V*
1982

Alejandro Otero
*Board 1*
1976
Automatically generated similarity matches created by Lineage:

**Gerhard Richter**
*Woods (2)*
2005

**Georges Braque**
*Studio V*
1949-50

**Henri Matisse**
*Woman on a High Stool (Germaine Raynal)*
Paris, quai Saint-Michel, early 1914

Noya Kohavi  
April 2019
Automated similarity matches created by Lineage:

Odilon Redon  
*Apache (Man on Horseback)*  
c. 1875

Jean Dubuffet  
*The Cow with the Subtle Nose*  
September 1954
Automated similarity matches created by Lineage:

Nicholas Africano
*The Scream*
1976

Tomás Maldonado
*Development of a Triangle*
1949

Noya Konavi
April 2019
Automated similarity matches created by Lineage:

László Moholy-Nagy
*EM 2 (Telephone Picture)*
1923

Rubén Núñez
*Articulated Arrhythmic*
1955
Automated similarity matches created by Lineage:

Eugene Berman
The Good Samaritan
1930

Alan Davie
Magic Box
1955
Automated similarity matches created by Lineage:

Mark Rothko
No. 5/No. 24
1948

Ivan Kliun
Suprematism
1916

Noya Kohavi

André Masson
Card Trick
1923

April 2019
Automated similarity matches created by Lineage:

Gabriel Orozco
*Kytes Tree*
2005

Theo van Doesburg (Christian Emil Marie Küpper)
*Rhythm of a Russian Dance*
June 1918
Automated similarity matches created by Lineage:

Beauford Delaney
Composition 16
1954-56

Richard Diebenkorn
Berkeley, 46
1955
Automated similarity matches created by Lineage:

Anselm Kiefer  
*Ways of Worldly Wisdom: Arminius’s Battle*  
1978

Joaquin Torres-Garcia  
*Construction in White and Black*  
1938
Automated similarity matches created by Lineage:

William Baziotes
*Pompeii*
1955

Tony Smith
*Untitled*
1962, 1980

Noya Kohavi
April 2019
Automatically generated similarity matches created by Lineage:

Miroslav Sutej
_Bombardment of the Optic Nerve, II_
1963

Jackson Pollock
_Echo: Number 25, 1951_
1951

Noya Kohavi  
April 2019
Questions?