Connecting the Dots: Using Digital Scholarship Methods to Facilitate New Modes of Discovery in Special Collections

Carruthers, Matthew

http://hdl.handle.net/2027.42/138110
Creating the Future of Digital Scholarship Together: Collaboration from Within Your Library

Laurie Allen, University of Pennsylvania Libraries
Matt Carruthers, University of Michigan
Amy Hunsaker, University of Nevada, Reno
Dana Miller, University of Nevada, Reno

June 24, 2017

Co-sponsors:
LITA
ALCTS Cataloging and Metadata Management Section
ACRL Digital Humanities Interest Group
Using Digital Scholarship Methods to Facilitate New Modes of Discovery in Special Collections

Matt Carruthers
Metadata Projects Librarian
University of Michigan
Problem Statement

Distinct archival collections can often be very interrelated, containing materials by or about the same people, institutions, or families.

This interrelation can be key to a researcher’s understanding of the materials.

However, many current discovery interfaces for archival materials are not very good at representing the complex relationships between collections or entities represented in collections.
<table>
<thead>
<tr>
<th>Call Number / Description</th>
<th>Status</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>V+48 Browne Works</td>
<td>Building use only</td>
<td>Special Collections Archival and Manuscript Collections</td>
</tr>
<tr>
<td>Lepelade Labadie</td>
<td>Building use only</td>
<td>Special Collections Archival and Manuscript Collections</td>
</tr>
<tr>
<td>Am. Lt. Authors</td>
<td>Building use only</td>
<td>Special Collections Archival and Manuscript Collections</td>
</tr>
<tr>
<td>Access to the online version</td>
<td>Finding aid available online</td>
<td>Special Collections Archival and Manuscript Collections</td>
</tr>
</tbody>
</table>
Anne Waldman papers. Audio series. Part 2: WALDCA.31 - WALDCA.60
by Waldman, Anne, 1945-
Published: 20th century (exact date unknown)
- Audio (spoken word)
- Audio
- Archive

Location
Electronic Resources
Special Collections
Status
Finding aid available online
See holdings
Call Number / Description
Access to the online version
WALDCA.31 - WALDCA.60

Anne Waldman papers. Audio series. Part 3: WALDCA.61 - WALDCA.90
by Waldman, Anne, 1945-
Published: 20th century (exact date unknown)
- Audio (spoken word)
- Audio
- Archive

Location
Electronic Resources
Special Collections
Status
Finding aid available online
See holdings
Call Number / Description
Access to the online version
WALDCA.61 - WALDCA.90

Anne Waldman papers. Audio series. Part 4: WALDCA.91 - WALDCA.120
by Waldman, Anne, 1945-
Published: 20th century (exact date unknown)
- Audio (spoken word)
- Audio
- Archive

Location
Electronic Resources
Special Collections
Status
Finding aid available online
See holdings
Call Number / Description
Access to the online version
WALDCA.91 - WALDCA.120

Anne Waldman papers. Audio series. Part 5: WALDCA.121 - WALDCA.150
by Waldman, Anne, 1945-
Published: 20th century (exact date unknown)
- Audio (spoken word)
- Audio
- Archive

Location
Electronic Resources
Special Collections
Status
Finding aid available online
See holdings
Call Number / Description
Access to the online version
WALDCA.121 - WALDCA.150
Problem Statement Pt. 2

Relationships among archival collections and entities represented in collections are often not linear.

There is often a high level of interconnectedness and context that can be lost in traditional discovery environments.
Special Collections Digital Scholarship Team

Team:
Associate Director of Special Collections
Information Resources Specialist, Special Collections
Digital Scholarship Librarian
Metadata Projects Librarian

Goals:
1. Explore using digital scholarship methods and tools to enhance user experience and resource discovery in Special Collections.

2. Develop a pilot project as a proof of concept for a new service.
Result

Provide visualization on demand, customized to a patron’s individual research question.

- Use data from finding aids and publicly accessible databases along with open-source tools to create interactive visualizations of social networks of the entities and resources in our archival collections.
- Make connections between entities more apparent and provide a discovery tool to spur further inquiry and research.

ALA Annual 2017
June 24, 2017
Data Processing

Encoded Archival Description (EAD)
- Metadata standard for the encoding of finding aids for use in a networked (online) environment.

Encoded Archival Context – Corporate Bodies, Persons, and Families (EAC-CPF)
- Companion standard to Encoded Archival Description (EAD).
- Describes the entity responsible for creating (or who is represented in) one or more archival collections.
Data Processing: Tools

EAC-CPF record creation:

Social Networks and Archival Context
http://socialarchive.iath.virginia.edu/snac/search

Remixing Archival Metadata Project
http://rampeditor.info

xEAC
https://github.com/ewg118/xEAC
More Data Processing

Use XSLT to extract information about connections and format output files as tab-delimited text.

EAC-CPF + XSLT = Tab-delimited text

ALA Annual 2017
June 24, 2017
More Data Processing

Information extracted:

- Name of connection
- Type of entity
- Type of connection
- Link to more information (when available)

```xml
<cpfRelation
  xlink:arcrole="http://socialarchive.iath.virginia.edu/control/term#associatedWith"
  xlink:href="http://n2t.net/ark:/99166/w6f48mrk"
  xlink:role="http://socialarchive.iath.virginia.edu/control/term#Person"
  xlink:type="simple" xmlns:xlink="http://www.w3.org/1999/xlink">
  <relationEntry>Zirlin, Larry.</relationEntry>
</cpfRelation>
```
More Data Processing

Information extracted:

- Name of connection
- Type of entity
- Type of connection
- Link to more information (when available)

```xml
<cpfRelation>
  xlink:arcrole="http://socialarchive.iath.virginia.edu/control/term#associatedWith"
  xlink:href="http://n2t.net/ark:/99166/w6f48mrk"
  xlink:role="http://socialarchive.iath.virginia.edu/control/term#Person"
  xlink:type="simple" xmlns:xlink="http://www.w3.org/1999/xlink">
    <relationEntry>Zirlin, Larry.</relationEntry>
  </cpfRelation>
```
More Data Processing

Information extracted:
- Name of connection
- Type of entity
- Type of connection
- Link to more information (when available)

Entity types:
- Person
- Corporate Body
- Family
- Archival Resource
- Bibliographic Resource

```xml
<cpfRelation xlink:arcrole="http://socialarchive.iath.virginia.edu/control(term#associatedWith)"
  xlink:href="http://n2t.net/ark/99166/w6f48mrk"
  xlink:role="http://socialarchive.iath.virginia.edu/control(term#Person)"
  xlink:type="simple" xmlns:xlink="http://www.w3.org/1999/xlink">
  <relationEntry>Zirlin, Larry.</relationEntry>
</cpfRelation>
```
The Process

Information extracted:

- Name of connection
- Type of entity
- Type of connection
- Link to more information (when available)

Connection types:

- Associated With
- Corresponded With
- Creator Of
- Referenced In

```xml
<cpfRelation>
  xlink:arcrole="http://socialarchive.iath.virginia.edu/control/term#associatedWith"
  xlink:href="http://n2t.net/ark:/99166/w6f48mrk"
  xlink:role="http://socialarchive.iath.virginia.edu/control/term#Person"
  xlink:type="simple" xmlns:xlink="http://www.w3.org/1999/xlink">
    <relationEntry>Zirlin, Larry.</relationEntry>
</cpfRelation>
```
More Data Processing

Information extracted:

• Name of connection
• Type of entity
• Type of connection
• Link to more information (when available)

<xpfRelation>
xlink:arcrole="http://socialarchive.lath.virginia.edu/control/term#associatedWith"
xlink:href="http://n2t.net/ark:/99166/w6f48mrk"
xlink:role="http://socialarchive.lath.virginia.edu/control/term#Person"
xlink:type="simple" xmlns:xlink="http://www.w3.org/1999/xlink">
<relationEntry>Zirlin, Larry.</relationEntry>
</cpfRelation>
Visualizing the Data

Cytoscape – http://cytoscape.org/

• Open source desktop application for visualizing complex networks.
## Visualizing the Data

### Edge Color (Unselected) Mapping

<table>
<thead>
<tr>
<th>Edge Color (Unselected)</th>
<th>interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>associatedWith</td>
</tr>
<tr>
<td></td>
<td>correspondedWith</td>
</tr>
<tr>
<td></td>
<td>creatorOf</td>
</tr>
<tr>
<td></td>
<td>referencedIn</td>
</tr>
</tbody>
</table>
Sharing the Visualizations

Share visualizations via web-publishing companion platform to Cytoscape:

CyNetShare

http://cynetshare.ucsd.edu/#/

Working example:

http://goo.gl/eHYLUi
Sharing the Visualizations

CyNetShare

An easy way to share your network visualizations

- Project Web Site
- Documentation
- Network Data URL
  - Enter Network File URL
- Visual Style URL (Optional)
  - Enter Style URL
  - Visualize!
Sharing the Visualizations

goo.gl/eHYLUi
Sharing the Visualizations
Results

• Developed a proof of concept for a new service.
• Produce customized interactive visualizations of archival collections in less than 10 minutes.
• Requires no investment of money, digital infrastructure, or support from library IT.
Next Steps

• Consider accessibility issues
Next Steps

- Consider accessibility issues
- Continue to streamline the workflow
Next Steps

• Consider accessibility issues
• Continue to streamline the workflow
• Perform user testing (Incorporate feedback into service model)
Next Steps

- Consider accessibility issues
- Continue to streamline the workflow
- Perform user testing (Incorporate feedback into service model)
- Explore more options for using and manipulating the data
Final thoughts

In general, the viability of a visualization or discovery service hinges on metadata.
Final thoughts

In general, the viability of a visualization or discovery service hinges on metadata.

So metadata processing and manipulation techniques will play a crucial role.
Final thoughts

In general, the viability of a visualization or discovery service hinges on metadata.

So metadata processing and manipulation techniques will play a crucial role.

But it’s also only one piece of the puzzle in creating a viable product.
Final thoughts

In general, the viability of a visualization or discovery service hinges on metadata.

So metadata processing and manipulation techniques will play a crucial role.

But it’s also only one piece of the puzzle in creating a viable product.

Use existing technologies and tools to help you whenever possible.
Thanks!

Matt Carruthers
mcarruth@umich.edu
@mattadata2