

2016-03-09

Making new discoveries from old data: Utilizing digital scholarship to foster new research in Special Collections

Carruthers, Matthew

<https://hdl.handle.net/2027.42/138108>

<http://creativecommons.org/licenses/by-nc-sa/4.0/>

Downloaded from Deep Blue, University of Michigan's institutional repository

Making new discoveries from old data

Utilizing digital scholarship to foster new research in
special collections

Slides: goo.gl/smvkvf

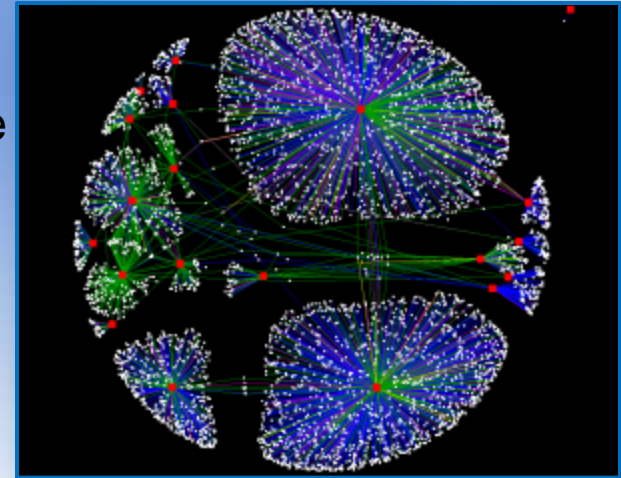
Goals

1. Explore using digital scholarship methods and tools to enhance user experience in special collections and potentially provide new services.
2. Develop a pilot project as a proof of concept for a new service.
 - Utilize existing open source technologies.
 - Little to no need for dedicated IT support.

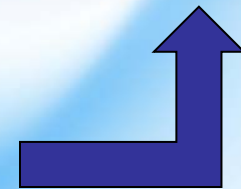
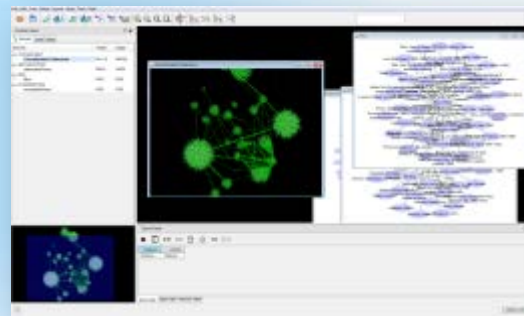
Result

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 creators of our archival collections.
- Make connections between entities more apparent and provide a discovery tool to spur further inquiry and research.



```
1 # Create a graph from the data
2 graph = nx.Graph()
3
4 # Add nodes
5 for creator in creators:
6     graph.add_node(creator)
7
8 # Add edges
9 for relationship in relationships:
10     graph.add_edge(relationship['creator1'], relationship['creator2'])
11
12 # Visualize the graph
13 nx.draw(graph, with_labels=True)
```



Data Processing

EAD

[illegible]

EAC-CPF

[illegible]

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

[illegible][illegible][illegible]

More Data Processing

Information extracted:

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

```
<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)

```
<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

```
<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

```
<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:

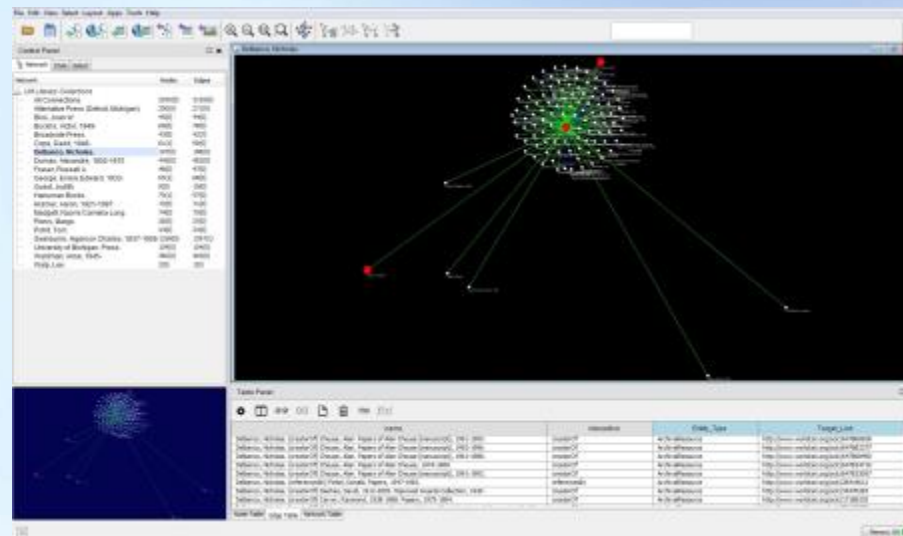
- Name of connection
- Type of entity
- Type of connection
- Link to more information (when available)

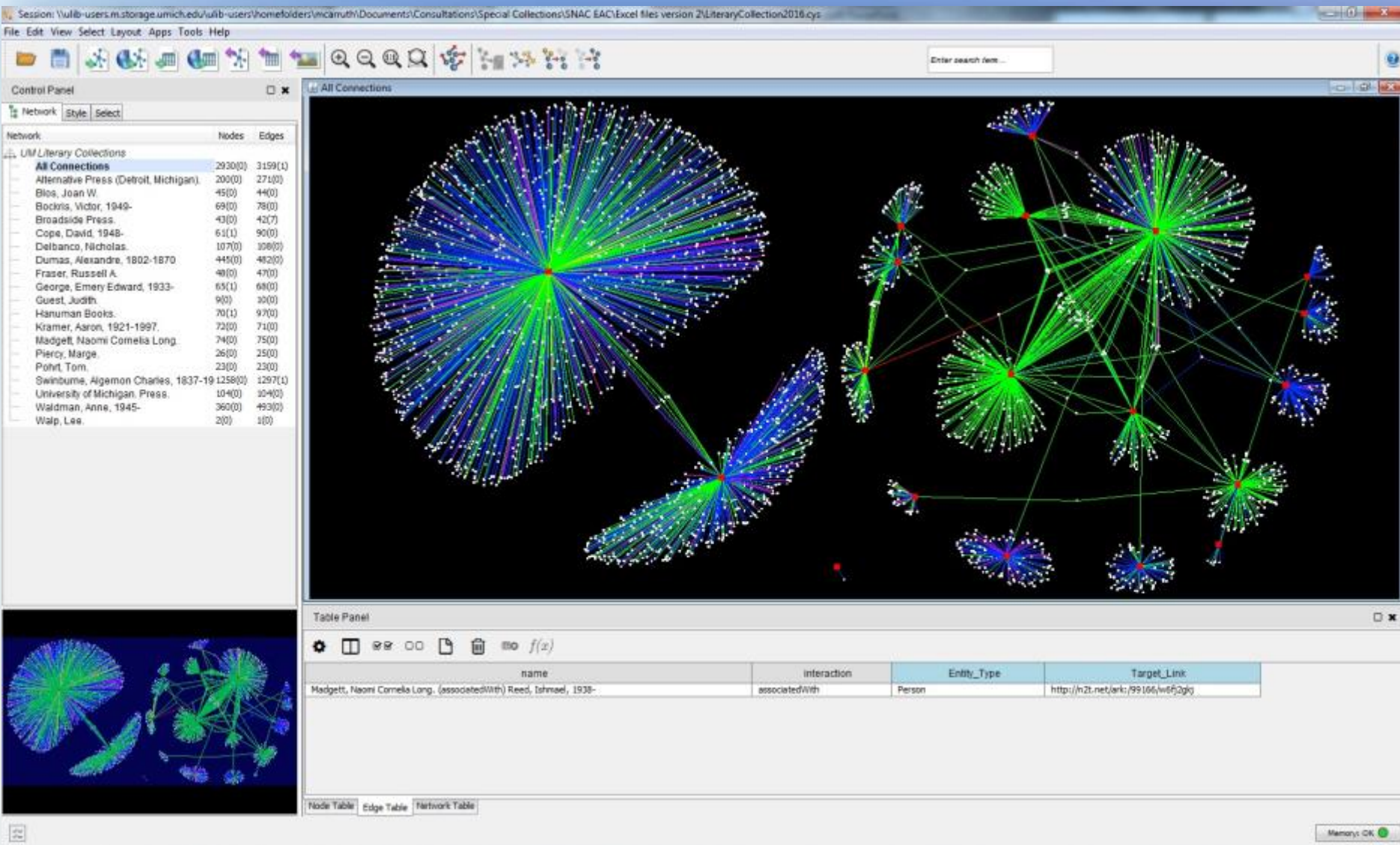
```
<cpfRelation
  xlink:arcrole="http://socialarchive.iath.virginia.edu/control/term#associatedWith"
  xlink:href="http://n2t.net/ark:/99166/w6f48mrk"
  xlink:role="http://socialarchive.iath.virginia.edu/control/term#Person"
  xlink:type="simple" xmlns:xlink="http://www.w3.org/1999/xlink">
  <relationEntry>Zirlin,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

Visualizing the Data

Cytoscape – <http://cytoscape.org/>





- Open source desktop application for visualizing complex networks.

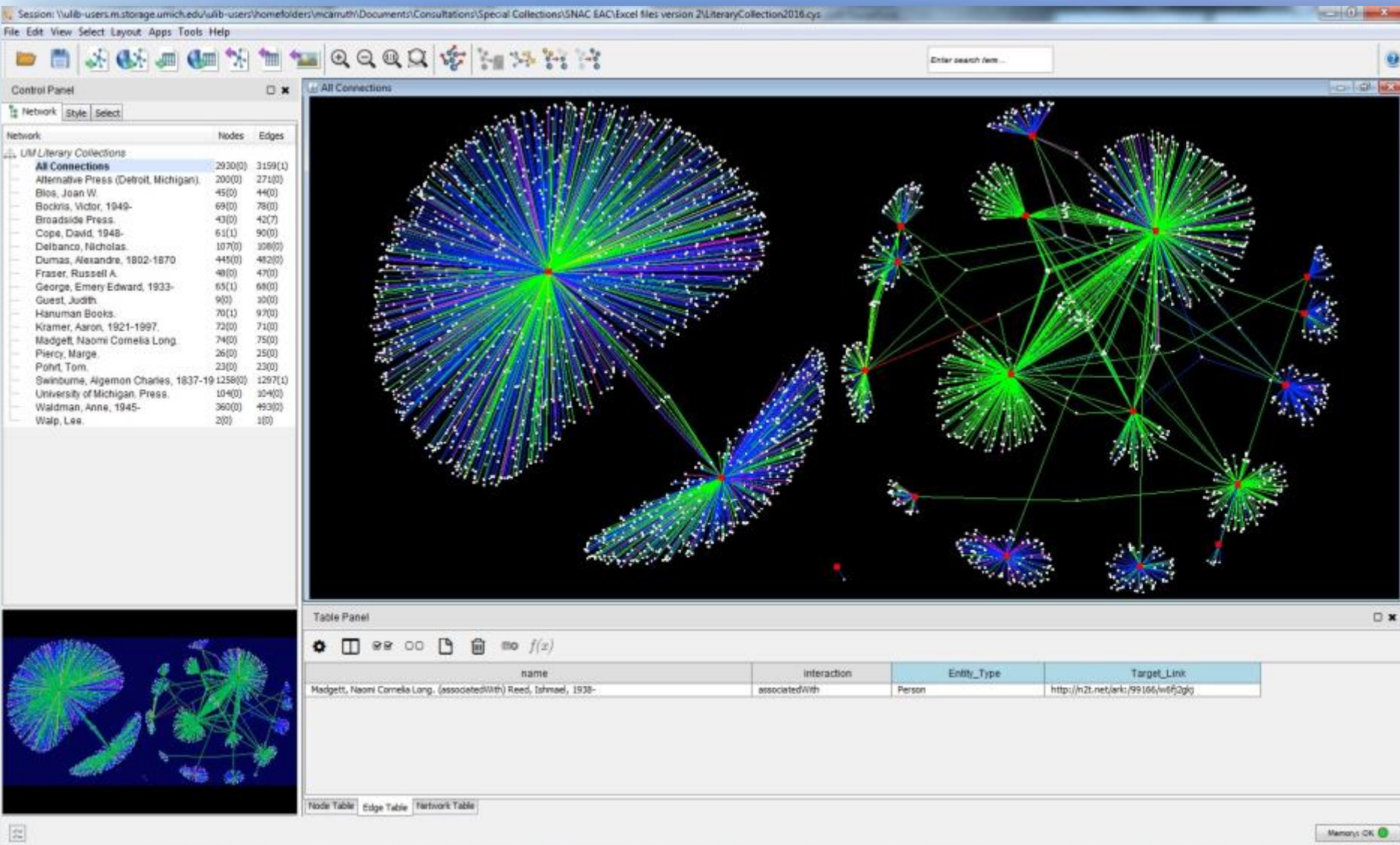




Visualizing the Data

Edge Color (Unselected) Mapping

Edge Color (Unselected)	interaction
	associatedWith
	correspondedWith
	creatorOf
	referencedIn





Enter search term...

Control Panel

Network	Nodes	Edges
UM Library Collections		
All Connections	2930(0)	3159(1)
Alternative Press (Detroit, Michigan)	200(0)	271(0)
Blos, Joan W.	45(0)	44(0)
Bockris, Victor, 1949-	69(0)	78(0)
Broadside Press	43(0)	42(7)
Cope, David, 1948-	61(1)	90(0)
Delbanco, Nicholas	207(0)	108(0)
Dumas, Alexandre, 1802-1870	445(0)	482(0)
Fraser, Russell A.	48(0)	47(0)
George, Emery Edward, 1933-	65(1)	68(0)
Guest, Judith	9(0)	10(0)
Hanuman Books	70(1)	97(0)
Kramer, Aaron, 1921-1997	72(0)	71(0)
Madgett, Naomi Cornelia Long	74(0)	75(0)
Piercy, Marge	26(0)	25(0)
Pont, Tom	23(0)	23(0)
Swinburne, Algernon Charles, 1837-19	1258(0)	1297(1)
University of Michigan Press	304(0)	104(0)
Waldman, Anne, 1945-	360(0)	493(0)
Walp, Lee	2(0)	1(0)

Delbanco, Nicholas

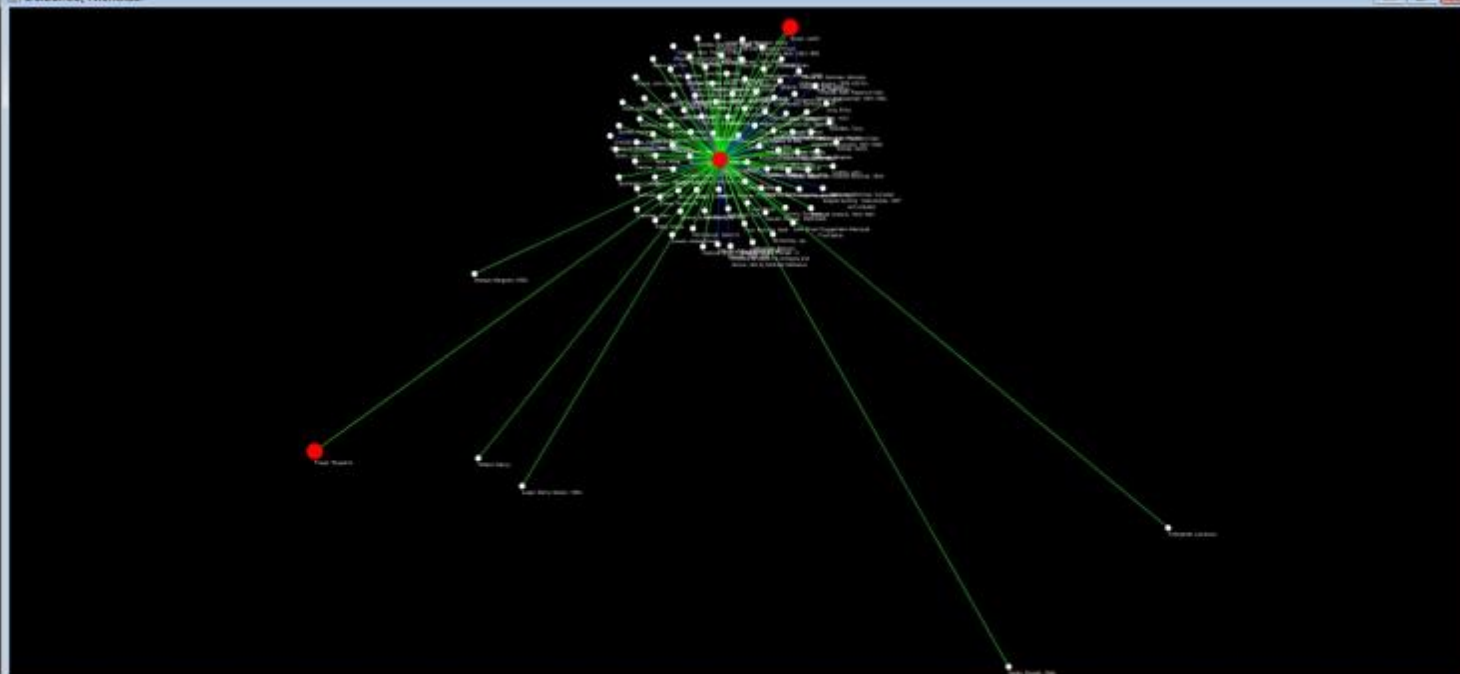
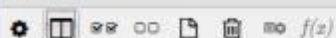


Table Panel



name	Interaction	Entity_Type	Target_Link
Delbanco, Nicholas. (creatorOf) Cheuse, Alan. Papers of Alan Cheuse [manuscript]. 1991-1995.	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/647860658
Delbanco, Nicholas. (creatorOf) Cheuse, Alan. Papers of Alan Cheuse [manuscript]. 1992-1996.	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/647883377
Delbanco, Nicholas. (creatorOf) Cheuse, Alan. Papers of Alan Cheuse [manuscript]. 1961-1988.	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/647808490
Delbanco, Nicholas. (creatorOf) Cheuse, Alan. Papers of Alan Cheuse [manuscript]. 1974-1989.	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/647824710
Delbanco, Nicholas. (creatorOf) Cheuse, Alan. Papers of Alan Cheuse [manuscript]. 1991-1992.	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/647833097
Delbanco, Nicholas. (referencedIn) Finkel, Donald. Papers, 1947-1985.	referencedIn	ArchiveResource	http://www.worldcat.org/oclc/28419612
Delbanco, Nicholas. (creatorOf) Gaches, David, 1912-2005. Hopwood Awards Collection, 1930-	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/34370265
Delbanco, Nicholas. (creatorOf) Carver, Raymond, 1938-1988. Papers, 1975-1984.	creatorOf	ArchiveResource	http://www.worldcat.org/oclc/17188335

Node Table Edge Table Network Table

Memory OK

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

Public Files

Gists

Network Data URL

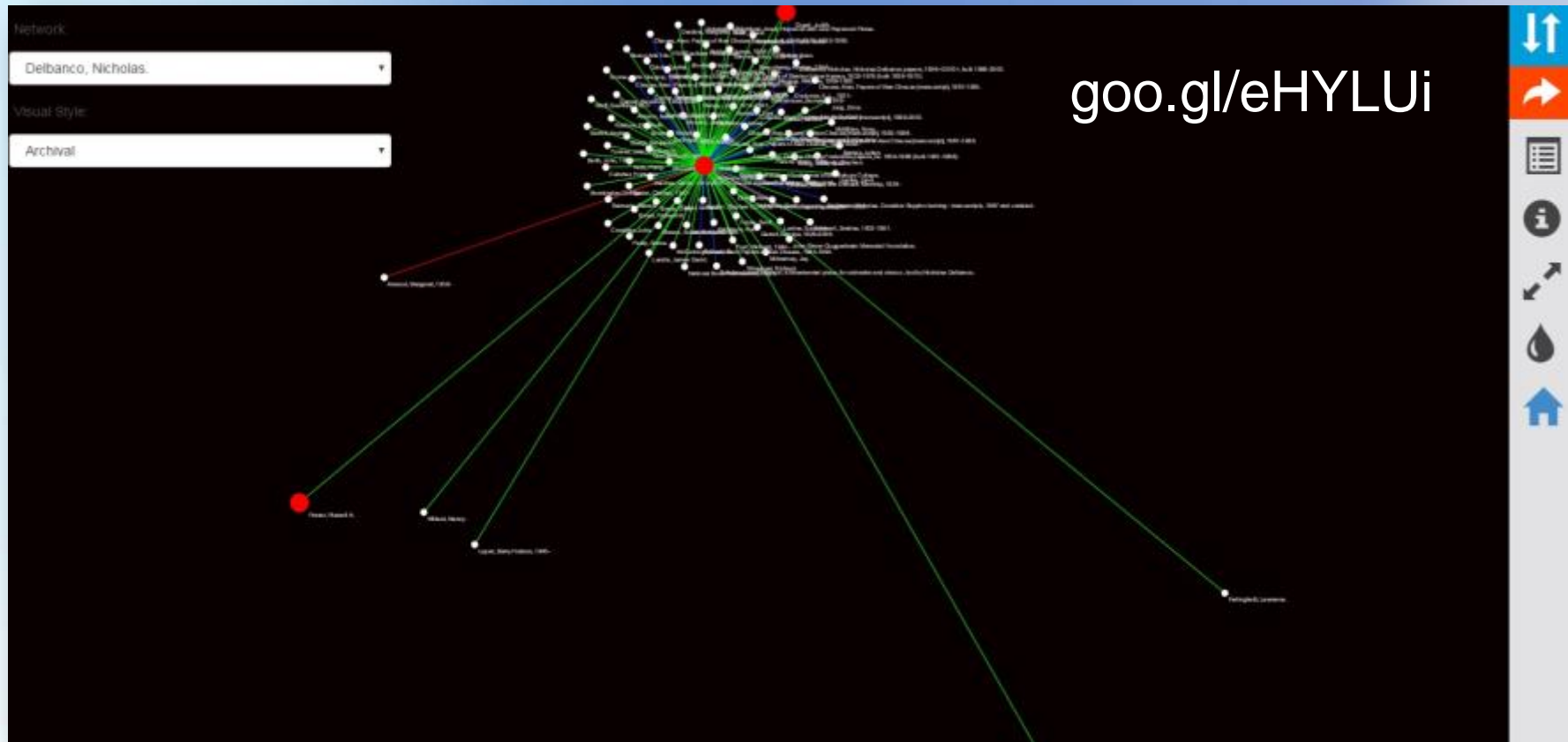
Enter Network File URL

Visual Style URL
(Optional)

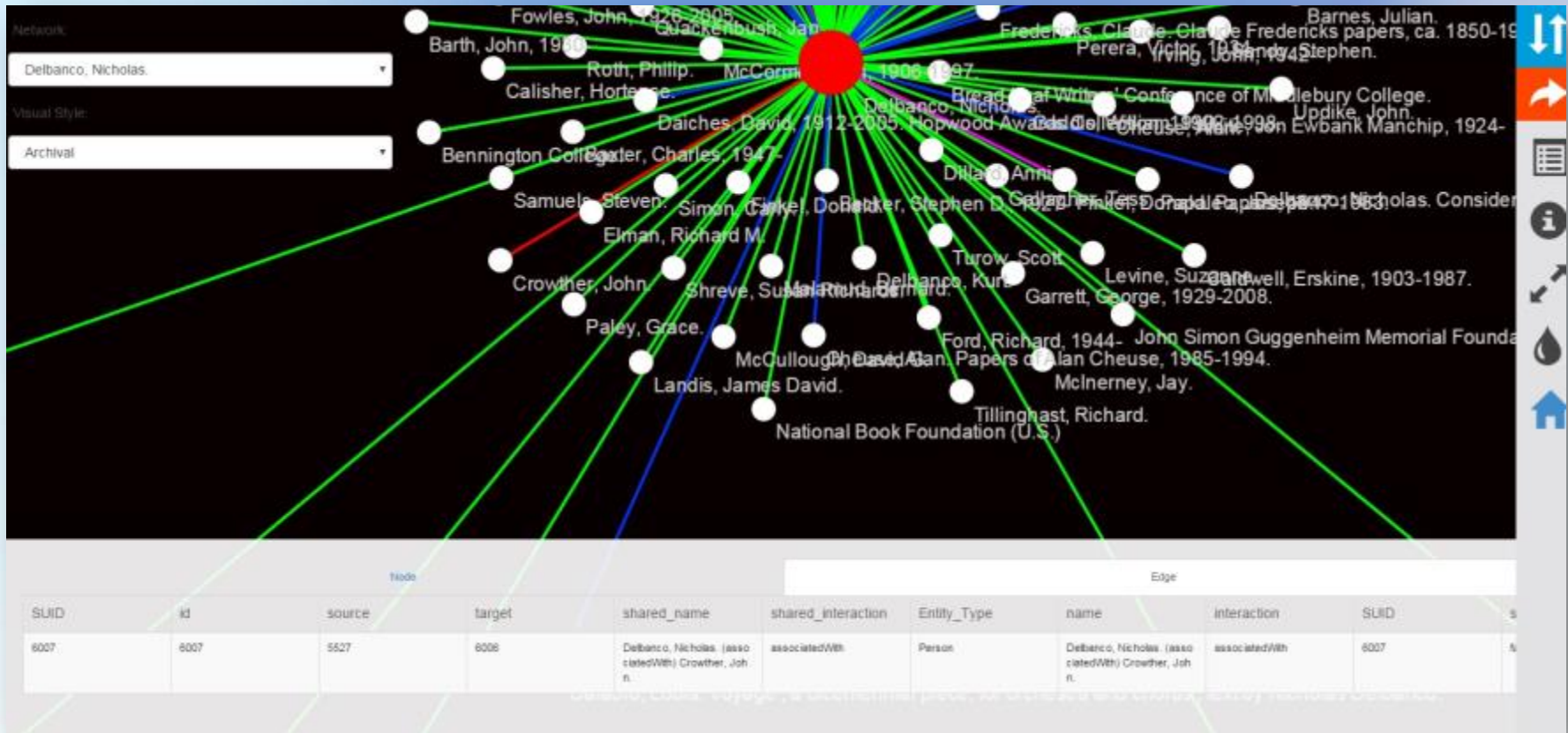
Enter Style URL

Visualize!

Sharing the Visualizations



Sharing the Visualizations



Next Steps

Continue to streamline the workflow

Next Steps

Perform user testing

(Incorporate feedback into service model)

Next Steps

Explore more options for using and
manipulating the data

(Customize and host local CyNetShare instance)

[http://www-
personal.umich.edu/~mcarruth/HanumanBooksVis/web_session/index.html#/](http://www-personal.umich.edu/~mcarruth/HanumanBooksVis/web_session/index.html#/)

Thanks!

Matt Carruthers

mcarruth@umich.edu

@mattadata2

Slides: goo.gl/smvkvf