## E Pluribus Unum: Melding Manifold Metadata for Use in a Consortial Setting

Tracy Seneca
University of Illinois at Chicago



Charles Blair University of Chicago

### We'll cover

- The CCC: who we are
- The current challenge for researchers
- The Chicago Portal: our vision
- Challenges for the CCC
- The technology selection
- Our approach: the Metadata Hopper

## The Chicago Collections Consortium



























## The Challenge for Researchers

#### John T. McCutcheon Collections



John T. McCutcheon, "There Ought to be School for Women Voters," 1913, Newberry Library Collection

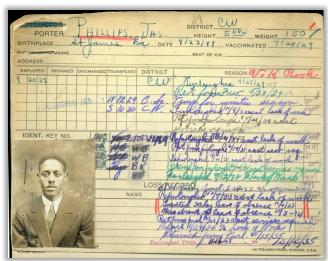
Chicago History Museum

**Newberry Library** 

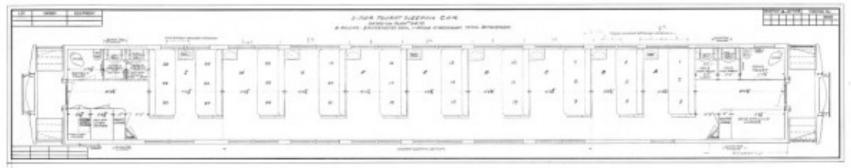
University of Illinois at Chicago

### **Pullman Collections**

Chicago Public Library
Newberry Library
Chicago History Museum
Art Institute of Chicago



Pullman, Employee Service Record, Phillips, Newberry Library Collection



An engineer drawing of a Pullman Company Three-Tier Tourist Sleeping Car, 1906, Newberry Library Pullman Digital Collection

## Century of Progress



Century of Progress - General Motors Building.
Art Institute of Chicago

## Libraries throughout Chicago ask:

How do we open up our archives and digital collections to Chicago researchers?

We did all this work to write finding aids, create collection records and digitize but nobody knows what's in our archives.

## Chicago Portal: The Vision

Depth of access

Chicago focus

Accessible to the novice but supports the seasoned researcher

Scheduled for August 2015 release Thanks to Andrew W. Mellon Foundation

## Chicago Portal: Scope

- ~ 1,500 EAD finding aids
- Hundreds of PDF finding aids with associated MARC
- ~ 2,000 collection level MARC records
- ~150,000 digital images
  - Custom CONTENTdm
  - MARC

## **Our Challenges**

## How do you express a name?

ead/archdesc/controlaccess/controlaccess/persname ead/archdesc/controlaccess/controlaccess/corpname ead/archdesc/controlaccess/controlaccess/famname ead/archdesc/did/origination/famname ead/archdesc/did/origination/corpname ead/archdesc/did/origination/persname ead/archdesc/descgrp/controlaccess/persname ead/archdesc/descgrp/controlaccess/corpname ead/archdesc/descgrp/controlaccess/famname ead/archdesc/controlaccess/corpname ead/archdesc/controlaccess/famname ead/archdesc/controlaccess/persname ead/archdesc/controlaccess/p/famname ead/archdesc/controlaccess/p/corpname ead/archdesc/controlaccess/p/persname ArchitectDesignerCreator AdditionalArchitectOrDesigner ProjectPersonnel Creator creator **Photographers** Photographer

# 14,730 subject terms (9,810 only ever used once)

```
Balloon framing
Burlesque (Theater) --
Burlesque (Theater –
Hack writers
Murgatroyd, Everett F., 1880 or 1-1946.
one way streets
Painting.
Paintings
Radio -- Illinois -- Chicago Metropolitan Area -- Receivers and reception -- 1980-1989. -- Ictgm
tree
Trees -- Chicago Metropolitan Area -- 1900-1919. – Ictgm
Weather experiments
Xylophone -- Instruction and study
```

#### Assume

 No one is going to adjust their local standards or practices

No one is going to maintain a second version of metadata for the CCC

#### Don't Assume

 That we can just point to CCC members' image records – some have digital images with no discovery platform

## **CCC Portal Technology Choice**

XTF (eXtensible Text Framework)

Flexible metadata processing

XTF allows variation

XTF allows full-text searching across all documents

A dc.xml "side-car" file accompanies each document

XTF works out of the box to index, search and display. Yet ...

XTF can easily be turned into a client/ server solution

Support

(http://xtf.cdlib.org)

#### **Test Data Set: Minimal Processing**

XTF has been optimized to recognize metadata element mapping, but none of Hopper's dc.xml files are included.

Search: Image in type [X]
Results: 881 Items
Sorted by: relevance 

Go!

Modify Search | New Search
Browse by Facet | Title | Author

Page: 1 2 3 4 5 ... Next

#### Subject

less 14th Street Yard (Chicago, III.) (14)Adler and Sullivan. (1) Adler, Dankmar, 1844-1900. (1)Advertisements. (1)(4)Aerial views African Americans -- Chicago --(1) Illinois African Americans -- Colorado --(1) Denver African Americans -- Illinois --(3)Chicago African Americans -- Nebraska --(1) Lincoln African-Americans -- Illinois --(2)Chicago

Air conditioning

Air-brakes

Alleys



Title: Borden Block

Description: Levy #690. Demolished 1917.

Date: 1881

Subjects: D. Adler & Co. | Adler and Sullivan. | Adler, Dankmar, 1844-1900. | Sullivan, Louis H., 1856-1924.

| Stores, Retail. | Commercial buildings. | Architecture--Illinois--Chicago. | Loop (Chicago, III.)



Title: Ellsworth Building

Description: [p. 9 of 163].

Date: 1986

Subjects: Van Osdel, John M., 1811-1891, I Van Osdel & Co. I Harry Weese and Associates, I Office

buildings, | Skyscrapers, | Architecture--Illinois--Chicago, | Loop (Chicago, III.)



(2)

(1)

(3)

Title: Ellsworth Building

Description: [p. 18 of 163].

Date: 1892

Subjects: Van Osdel, John M., 1811-1891. | Van Osdel & Co. | Harry Weese and Associates. | Young

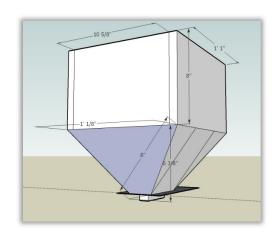
& Farrel Diamond Stone Sawing Co. | Office buildings. | Skyscrapers. | Architecture--

Illinois--Chicago. | Loop (Chicago, III.) | Advertisements.

| Coverage                                      |       |              |  |
|---|-------|--------------|--|
| less  |       | Title:       | Mildred B. Bevis Notebook of Designs for the Kalo Shop,  |
| <ul> <li>1514 S. Michigan Ave. [?]</li> </ul> | (1)   |              |  |
| • 1948  | (199) | Date:        |  |
| <ul> <li>2505 S. Michigan Ave.</li> </ul>     | (1)   |              | Art metal-workIllinoisChicago  |
| <ul> <li>2727 N. Hampden Ct.</li> </ul>       | (1)   | Subjects.    | Art metal-workillinoisChicago  |
| <ul> <li>39.755089, -104.988105</li> </ul>    | (9)   |              |  |
| <ul> <li>40.807701, -91.124793</li> </ul>     | (10)  | Title:       | Dankmar Adler (1844-1900) Collection,  |
| <ul> <li>40.813628, -96.70262</li> </ul>      | (5)   |              |  |
| <ul> <li>40.819008, -91.178</li> </ul>        | (21)  | Doddinpaoin  | Adler, partner of architect Louis H. Sullivan.   |
| <ul> <li>40.849678, -96.63443</li> </ul>      | (29)  | Date:        | 1941   |
| <ul> <li>40.936949, -94.896042</li> </ul>     | (6)   | Subjects:    | Chicago Auditorium Building (Chicago, III.)   ArchitectureIllinoisChicago19th century  |
| <ul> <li>40.947522, -90.37819</li> </ul>      | (12)  | •            | Sources   Architecture-IllinoisChicago20th centurySources   ArchitectsIllinoisChicago  |
| <ul> <li>41.061029, -94.364989</li> </ul>     | (6)   |              | Archives   |
| <ul> <li>41.259751, -95.942363</li> </ul>     | (2)   |              |  |
| <ul> <li>41.605955, -87.534554</li> </ul>     | (1)   | 7141-        | Feet on D. 175-co. (4000, 4000). Collection  |
| <ul> <li>41.615121, -87.510218</li> </ul>     | (1)   |              | Fazlur R. Khan (1929-1982) Collection,   |
| <ul> <li>41.617363, -87.513708</li> </ul>     | (1)   | Description: | This collection holds professional papers and writings and a small group of personal papers of the Bangladeshi structural engineer Fazlur R. Khan. |
| <ul> <li>41.619708, -87.519809</li> </ul>     | (1)   | Date:        |  |
| <ul> <li>41.620842, -87.52251</li> </ul>      | (1)   | Subjects:    | John Hancock Center (Chicago, III.)   Sears Tower (Chicago, III.)   SkyscrapersDesigns and   |
| <ul> <li>41.622195, -87.525006</li> </ul>     | (1)   | Subjects.    | plans   Structural engineering   Engineering design  |
| <ul> <li>41.739266, -87.541035</li> </ul>     | (1)   |              | F  |
| <ul> <li>41.759799, -87.827456</li> </ul>     | (1)   |              |  |
| • 41.76057, -88.320209                        | (23)  | Title:       | Peter J. Weber (1863-1923) Papers,   |
| • 41.772273, -87.619641                       | (1)   | Description: | Photographs, correspondence, and printed matter documenting a selection of German-   |
| • 41 777056 -87 580505                        | (1)   | -            | American architect Peter J. Weber's built and unbuilt projects in the United States and abroad,  |

## And...

XTF doesn't have an administrative interface

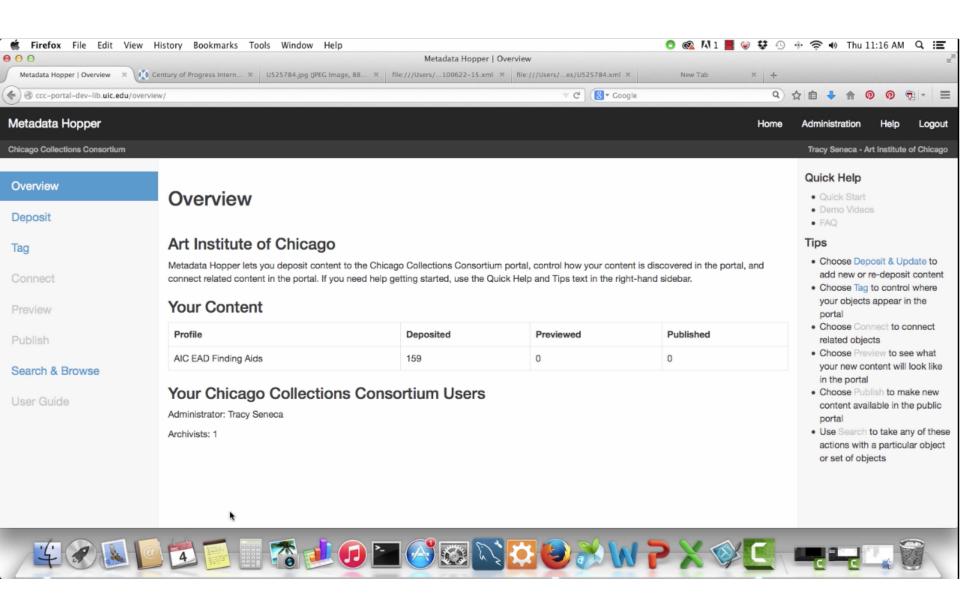


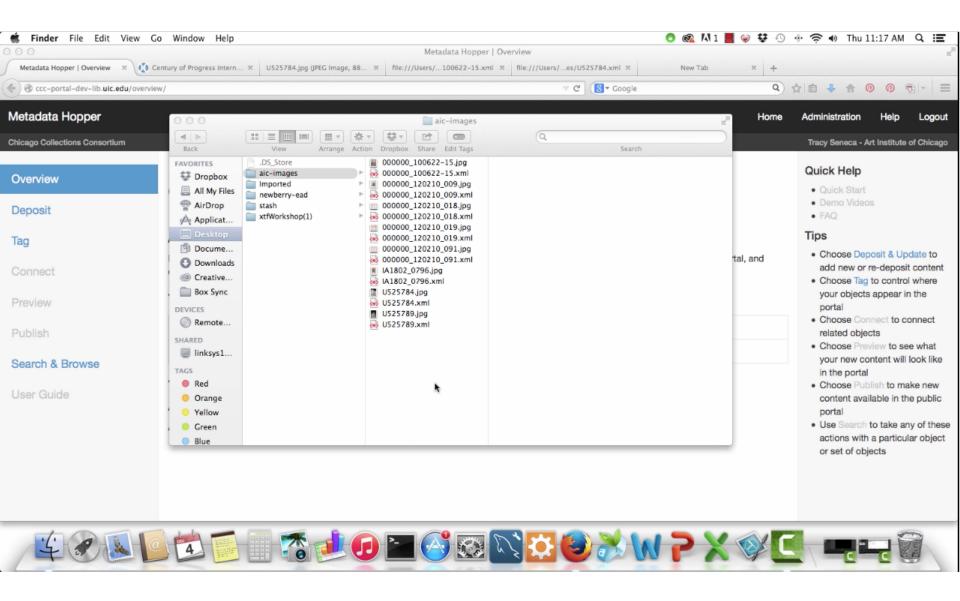
## The Metadata Hopper

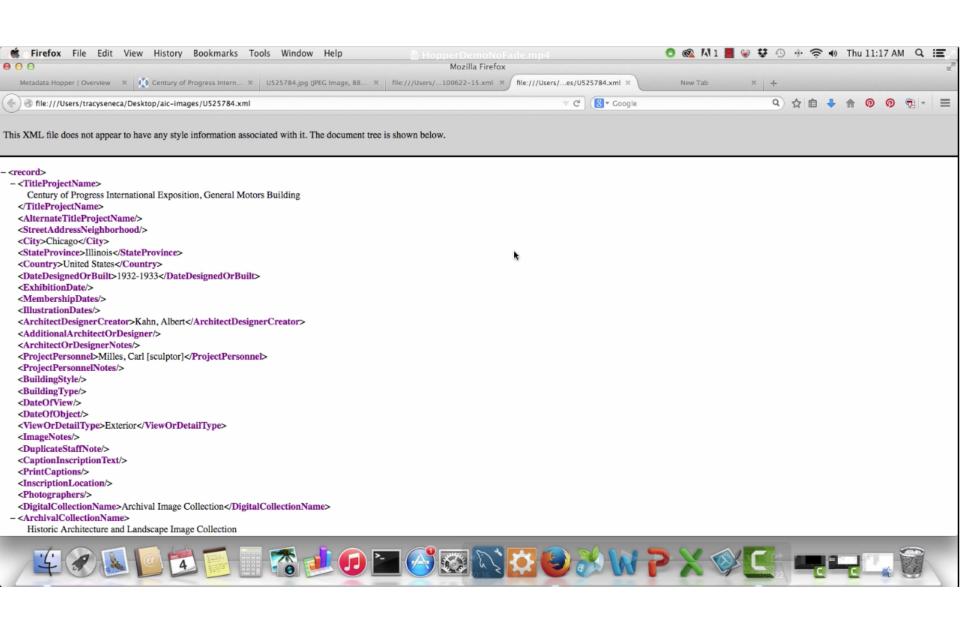
- Archivist/Librarian signs in
- Create a metadata map
- Add files
- Tag records (choose browsing facets)

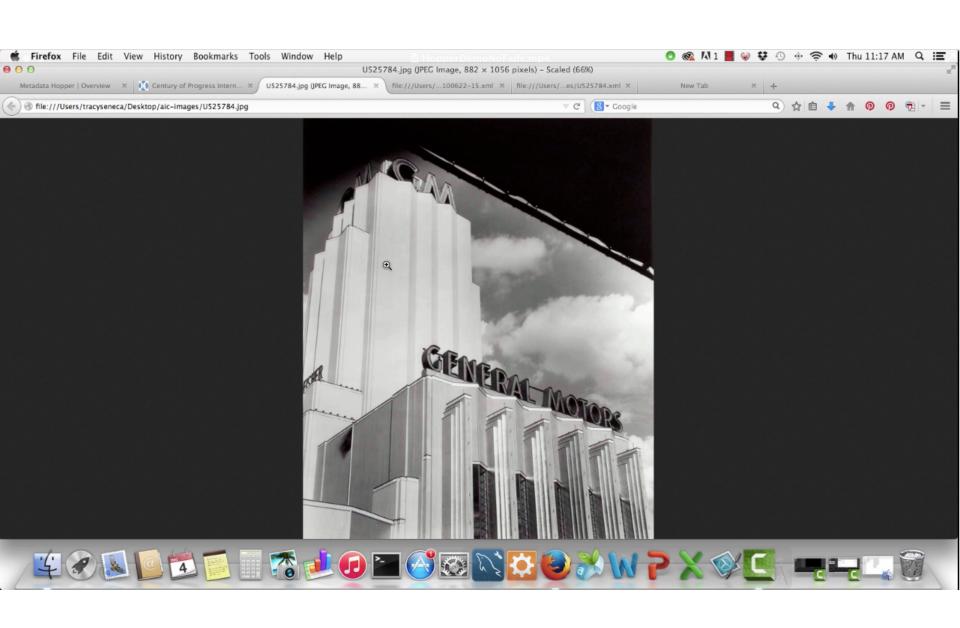
# Deposit: Create Import Rules (Metadata Map)

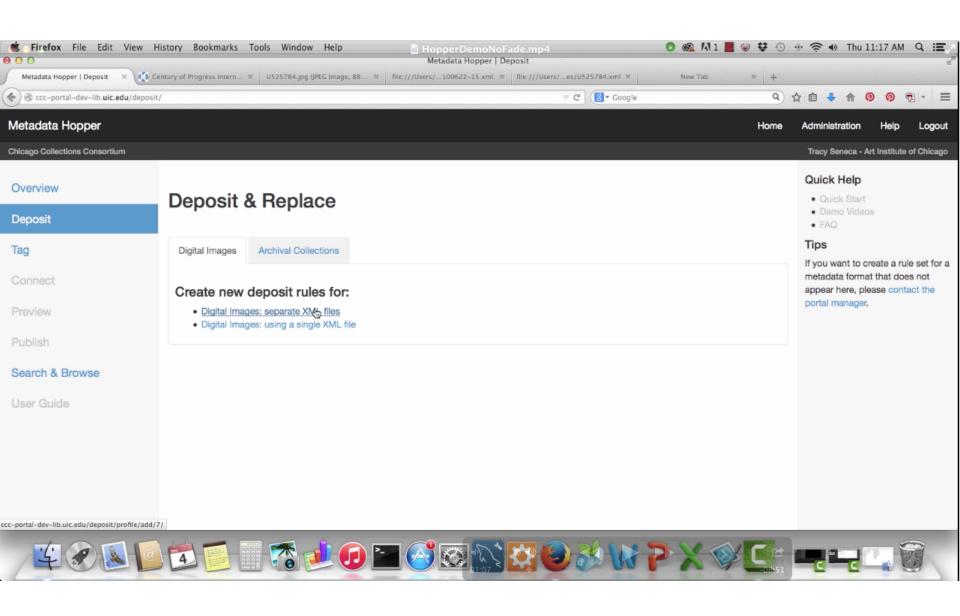
Art Institute of Chicago Digital Image Collection

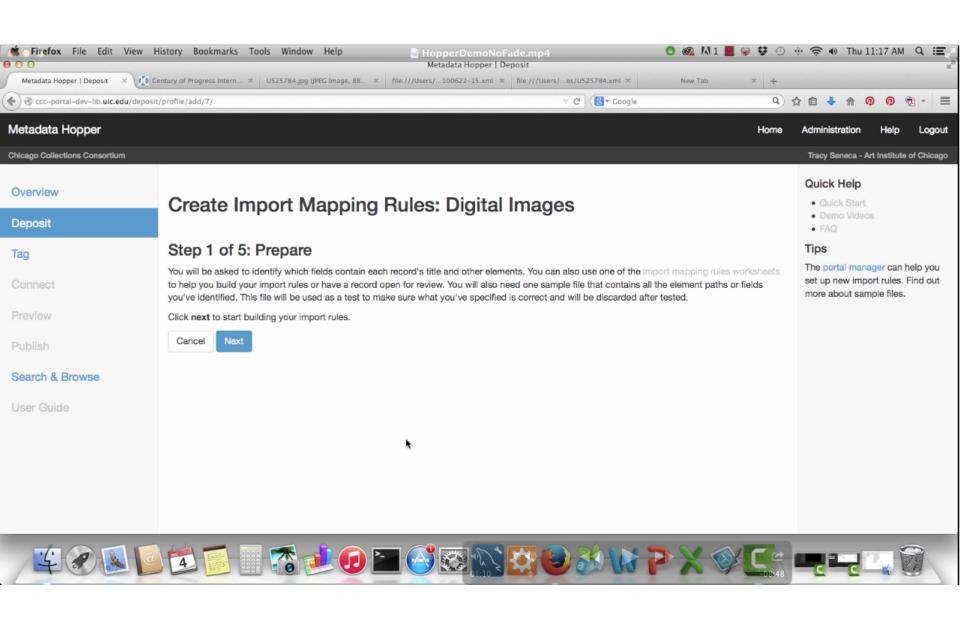


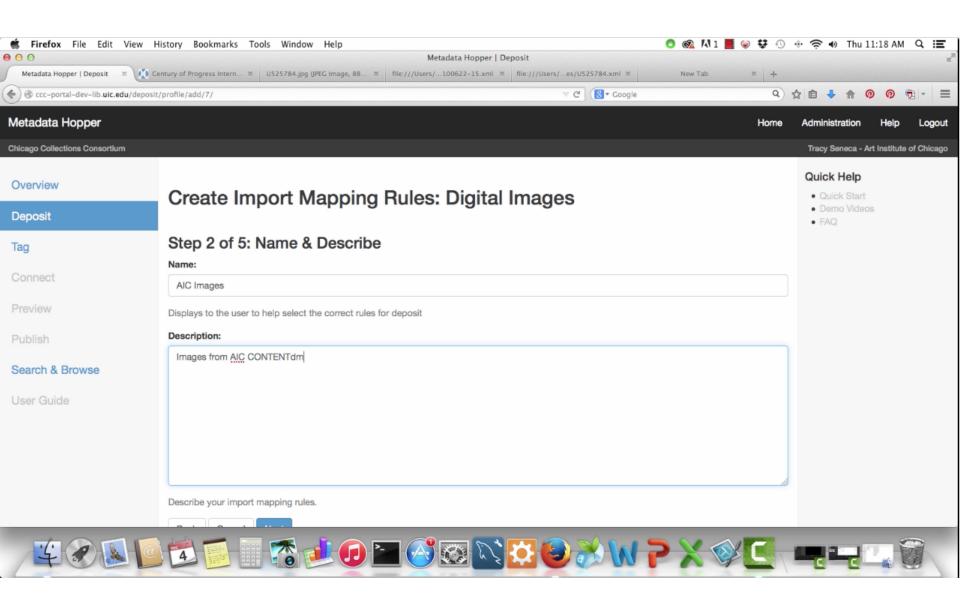


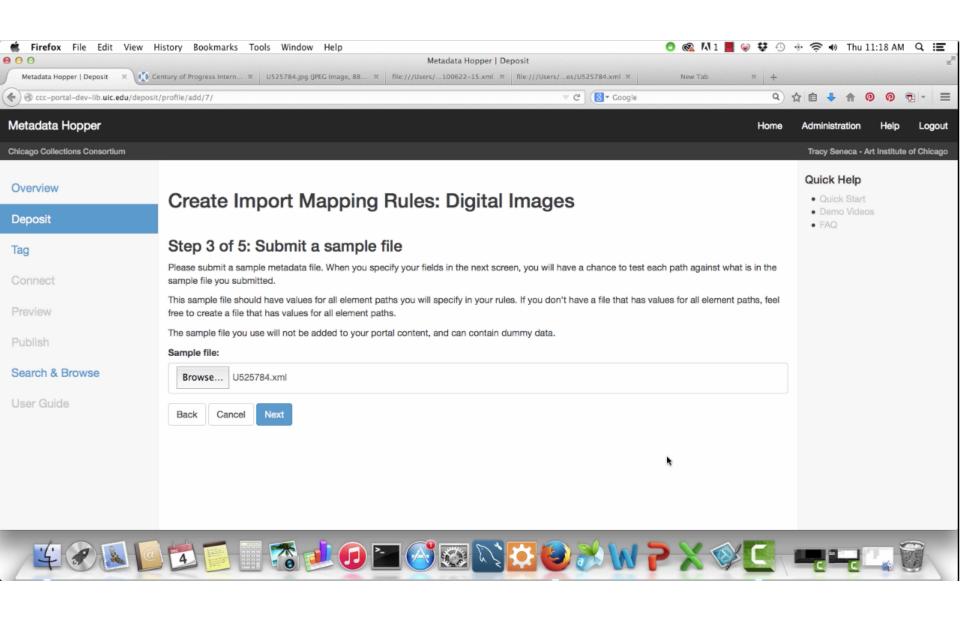


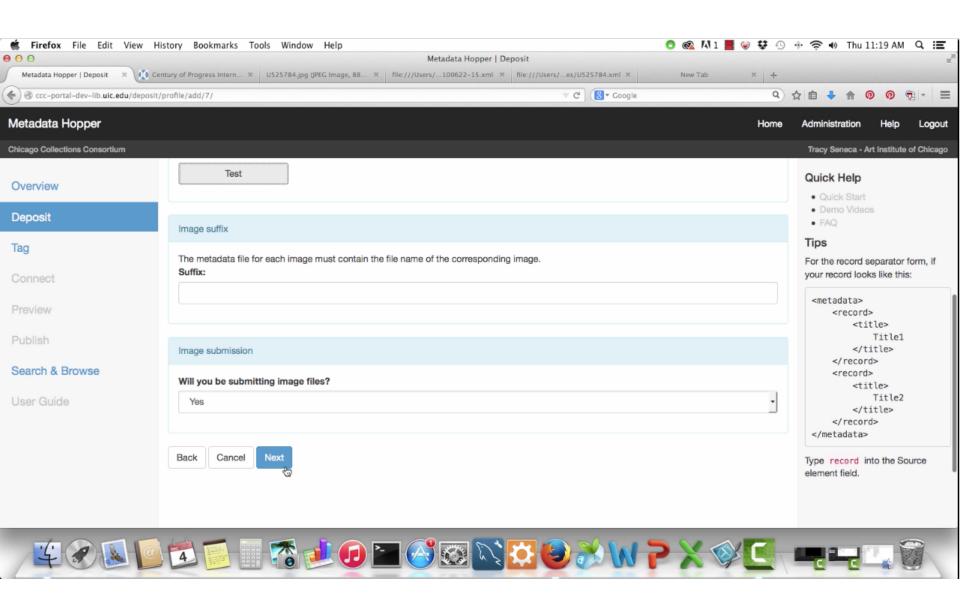


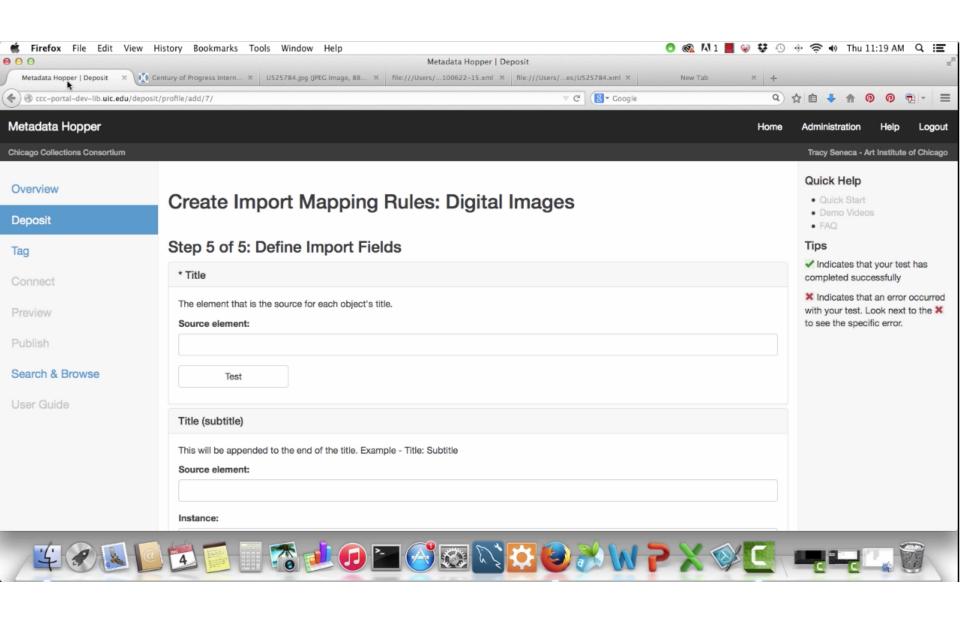


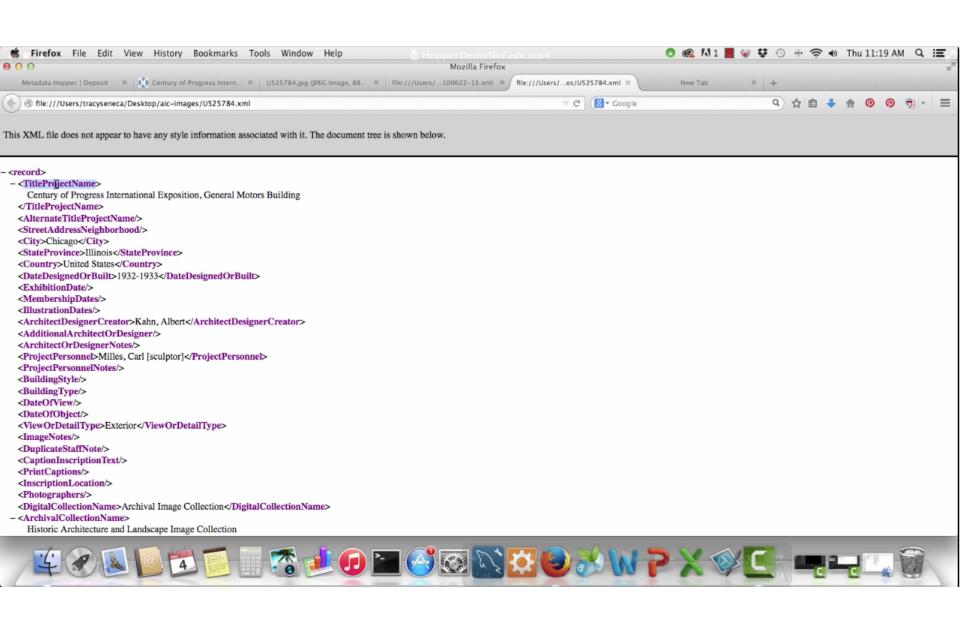


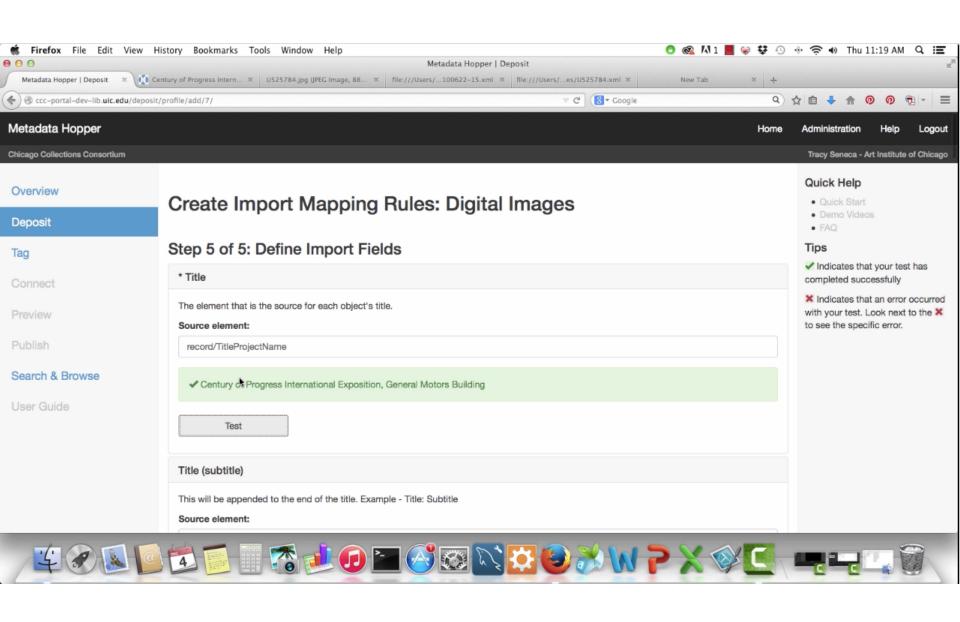


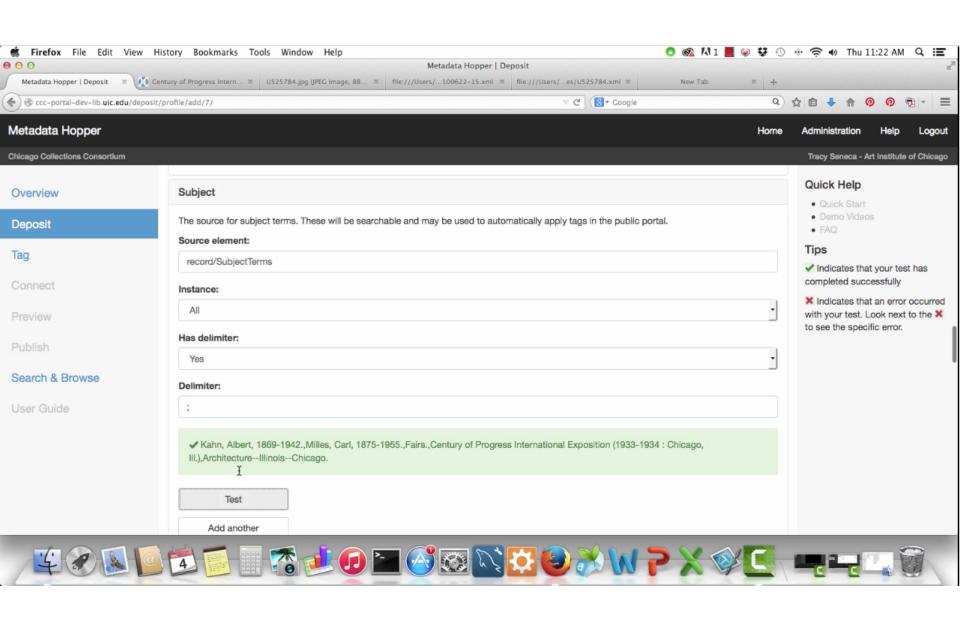






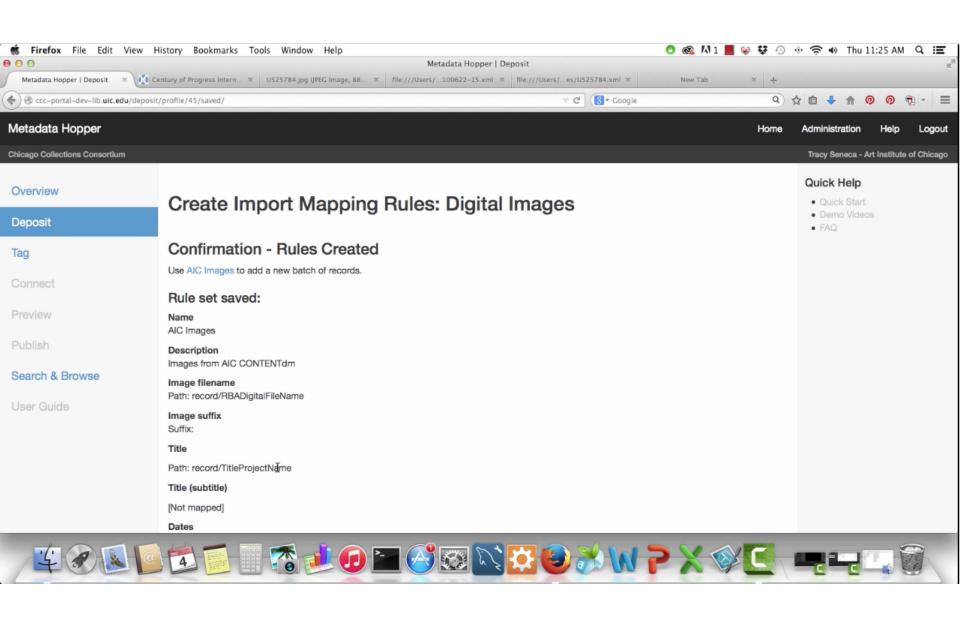






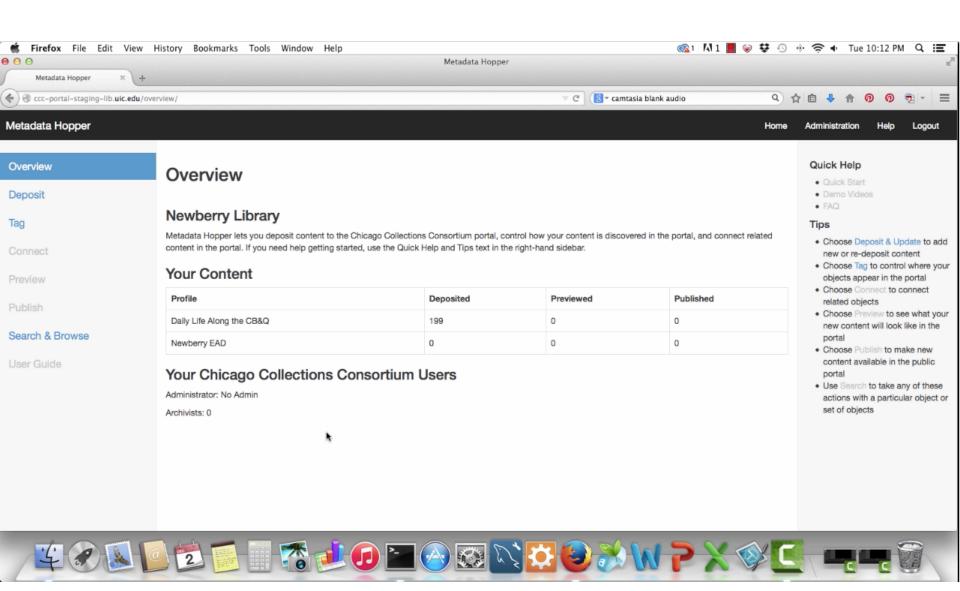
### We're mapping:

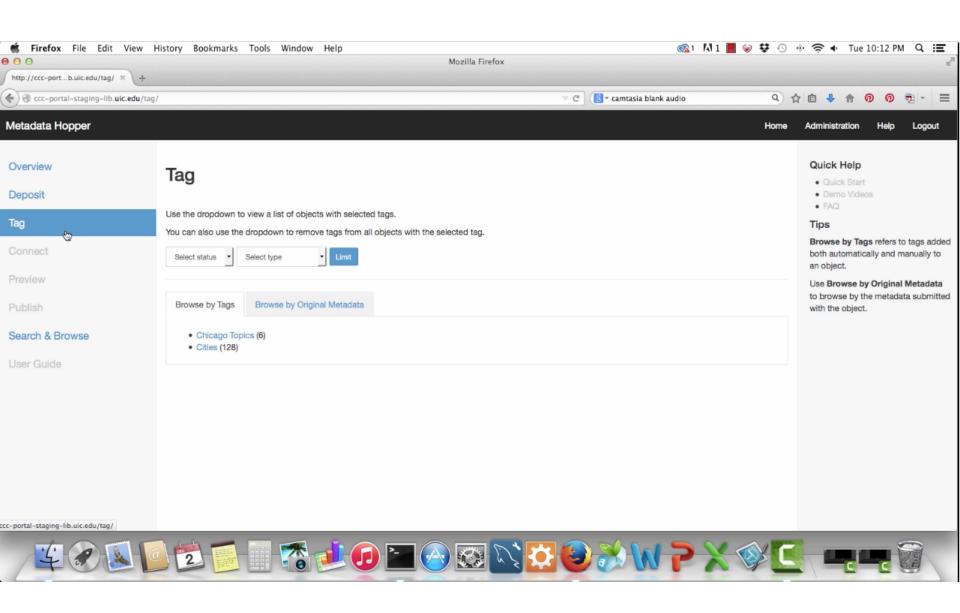
- Title (required)
- Subtitle
- Title (date for EAD)
- Dates
- Local Identifier
- Description
- Names
- Subject
- Geographic Coverage
- Source Archival Collection (for images)
- Digital Collection Name (for images)
- Digital Collection URL (for images)
- Rights

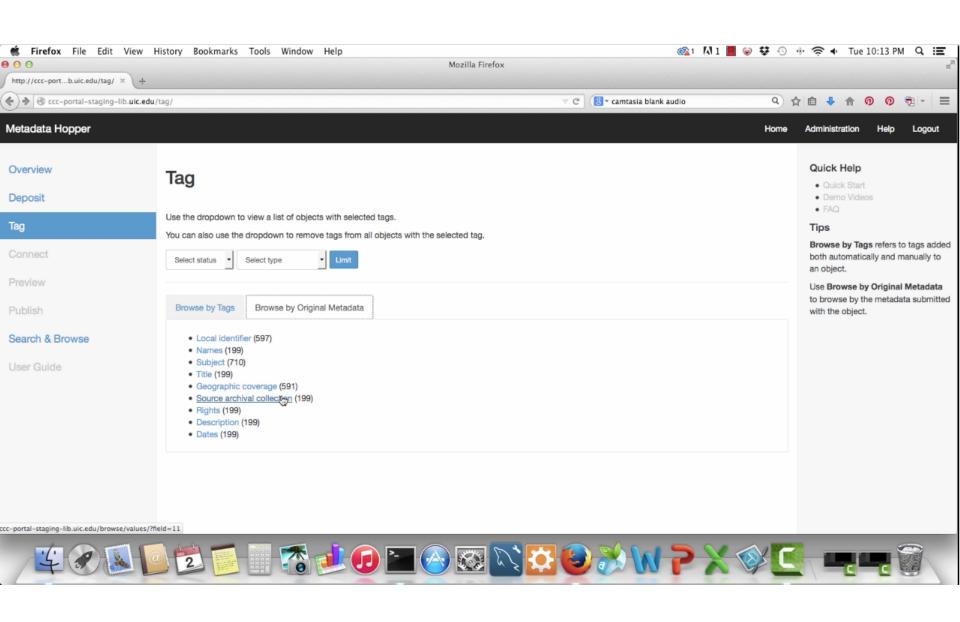


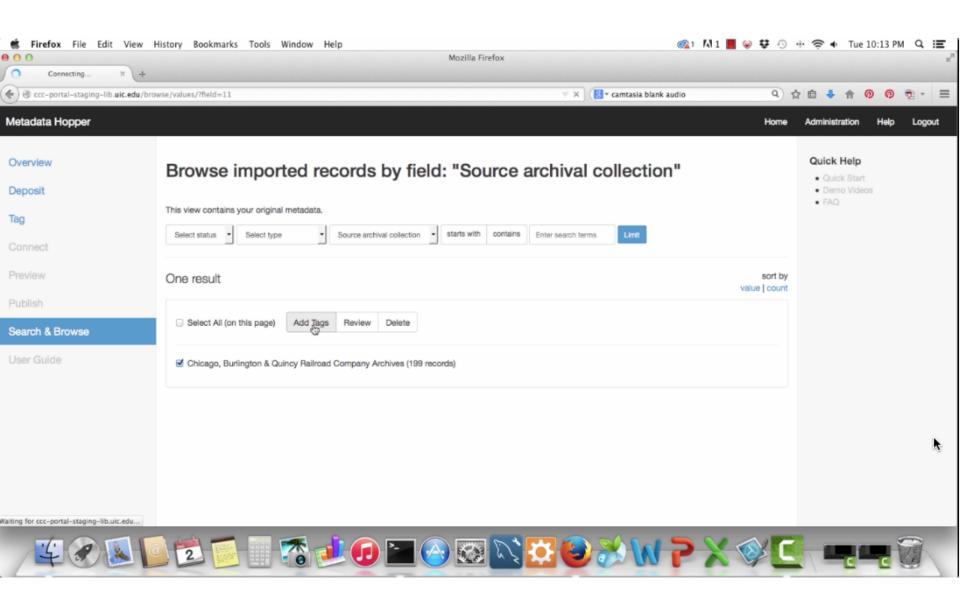
# Tagging: Simplify Topical Browsing

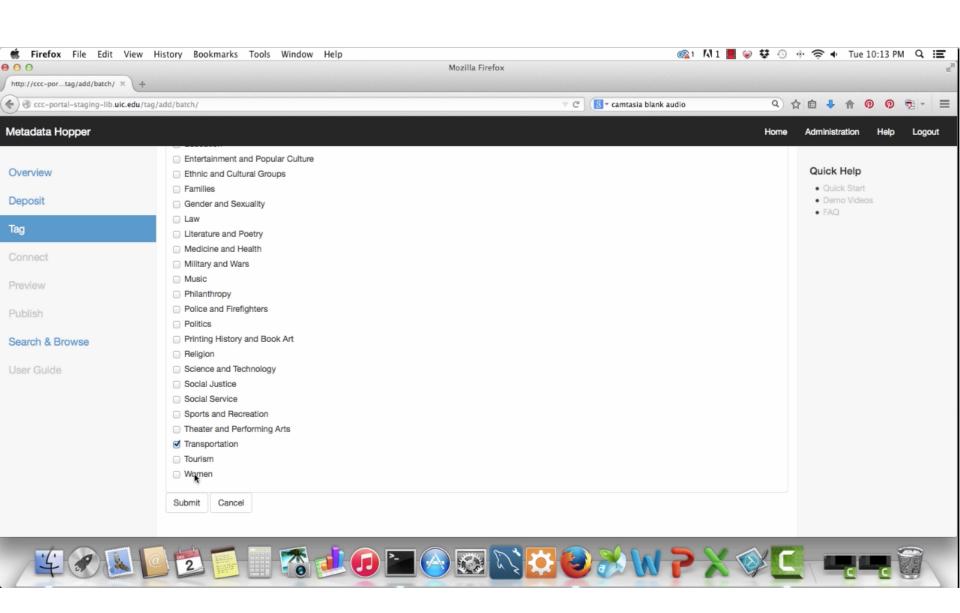
Newberry Library
Tag a Batch of Records

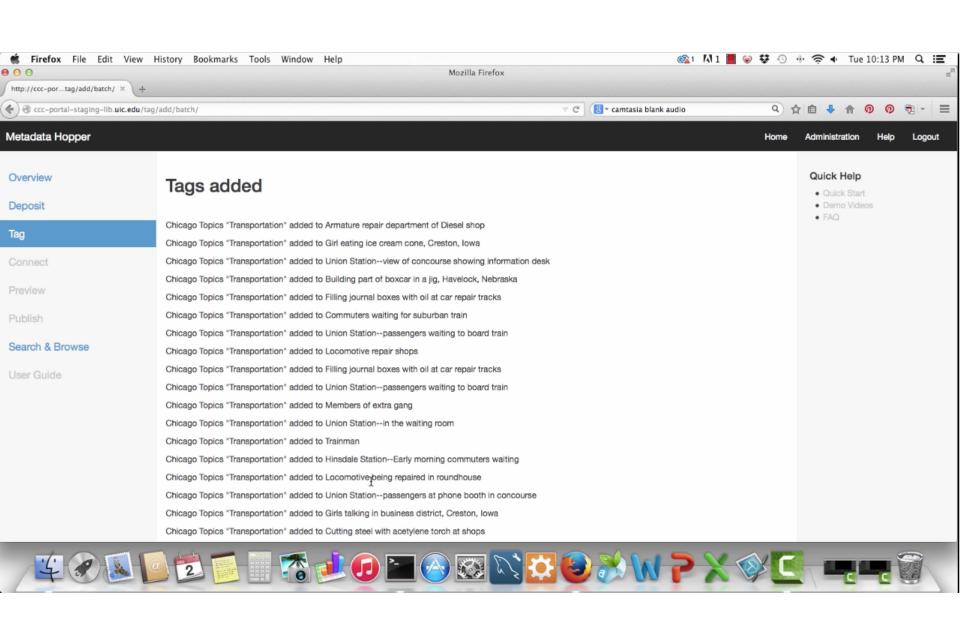


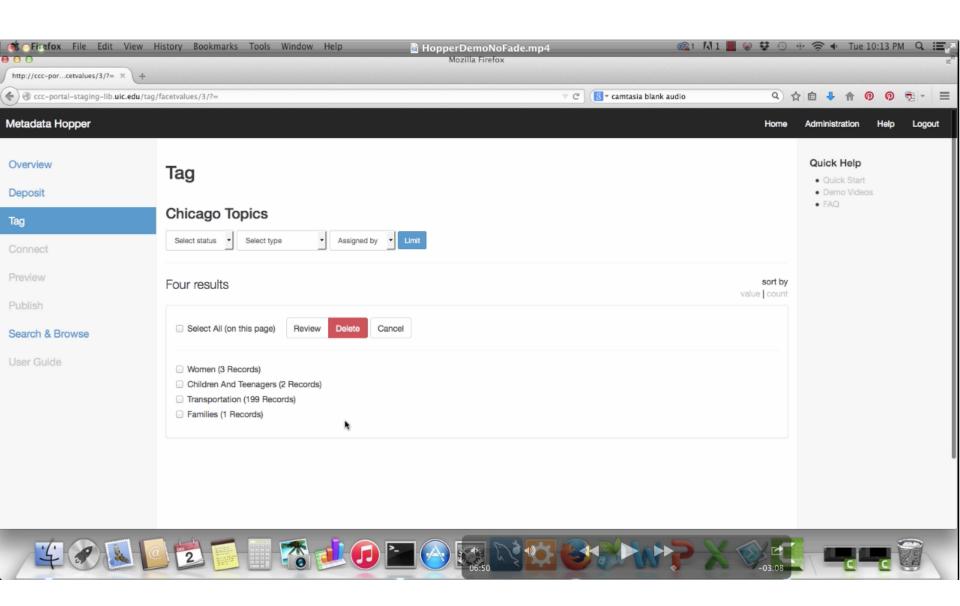






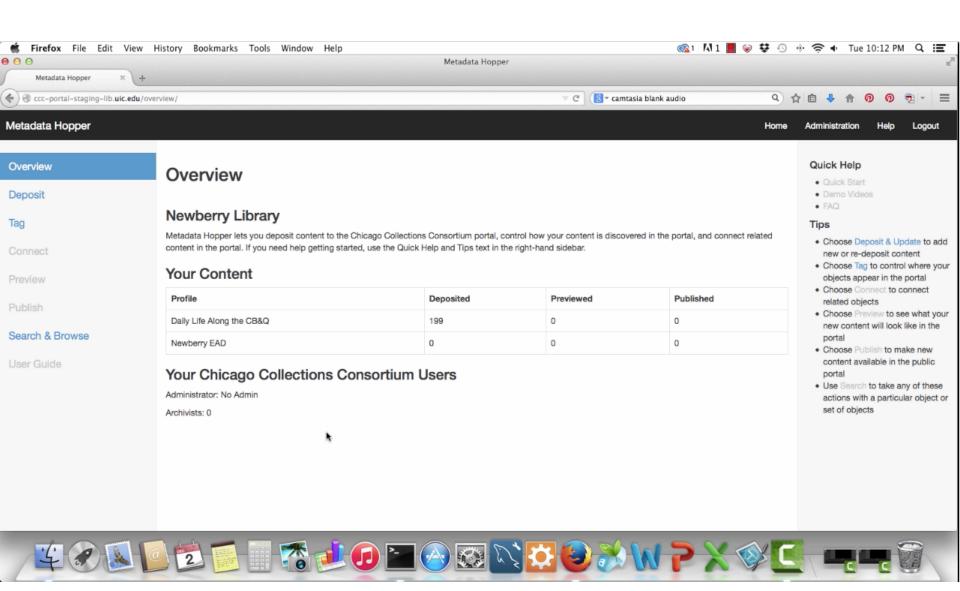


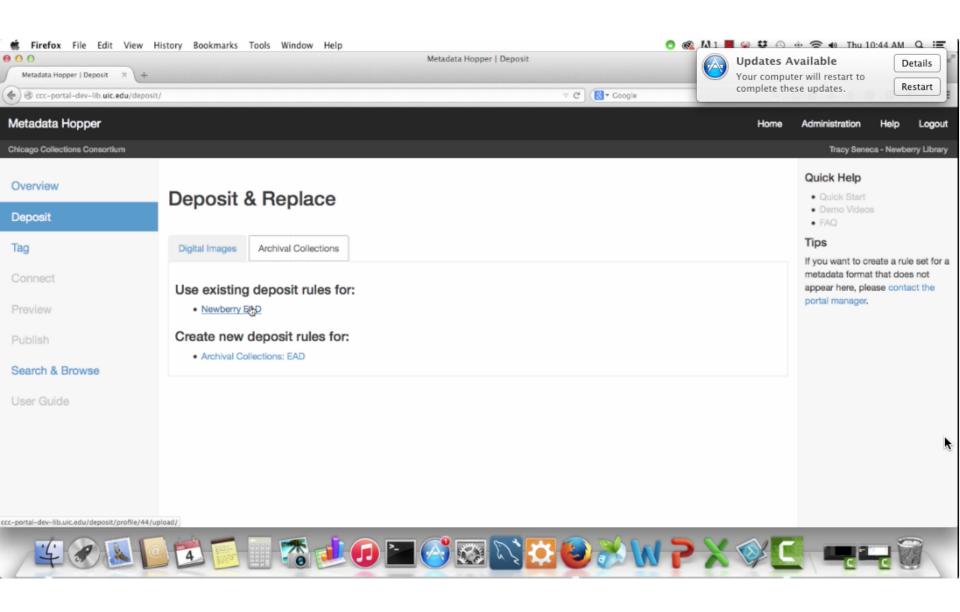


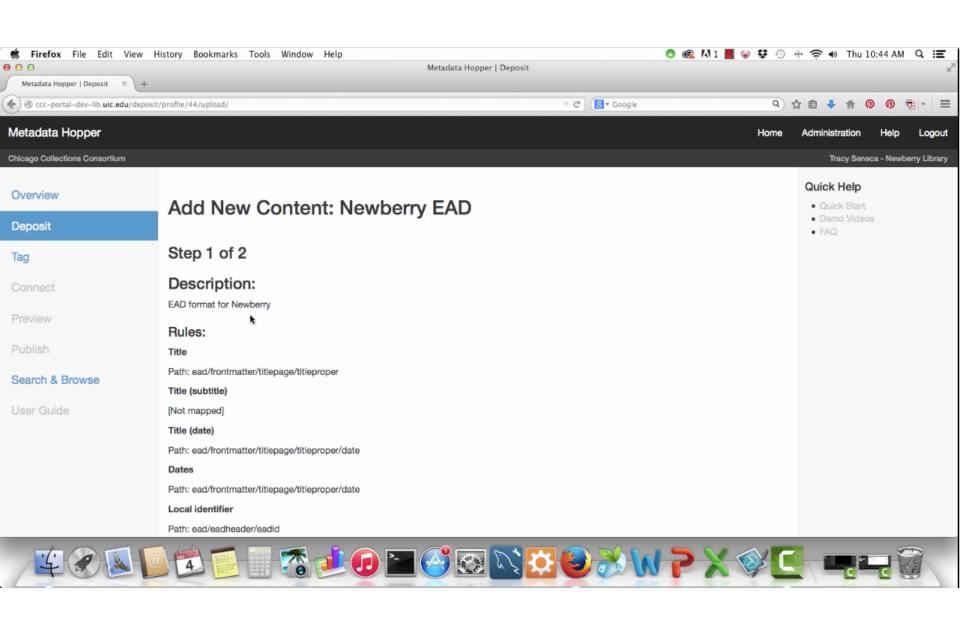


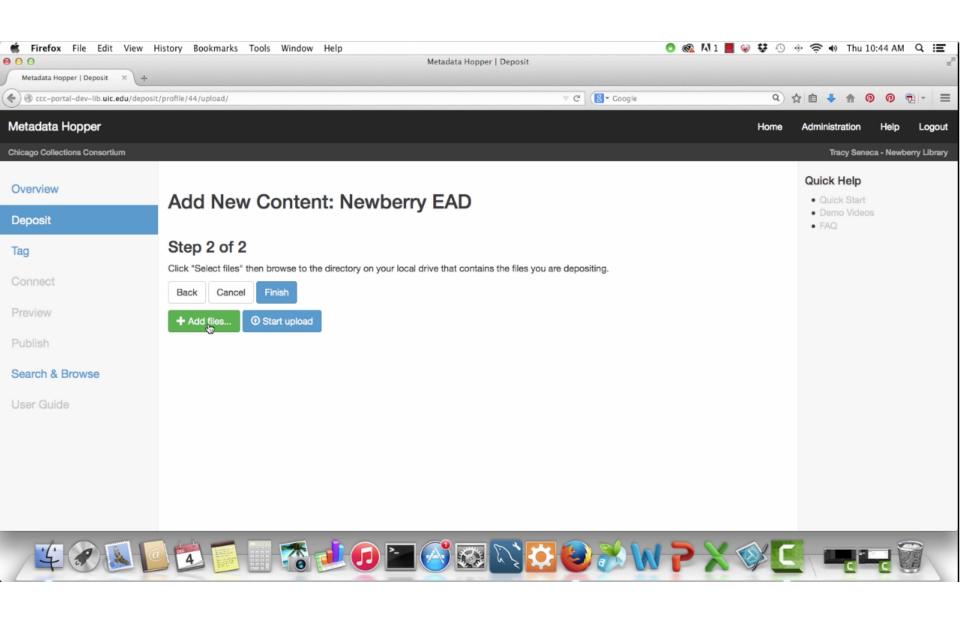
### Tagging: Autotag upon Deposit

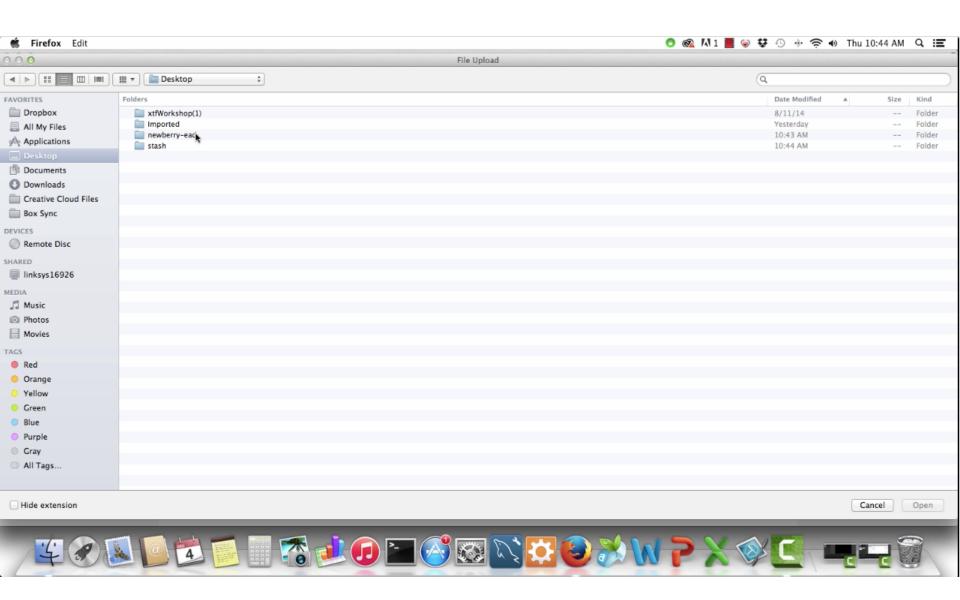
Newberry Library EAD Finding Aids

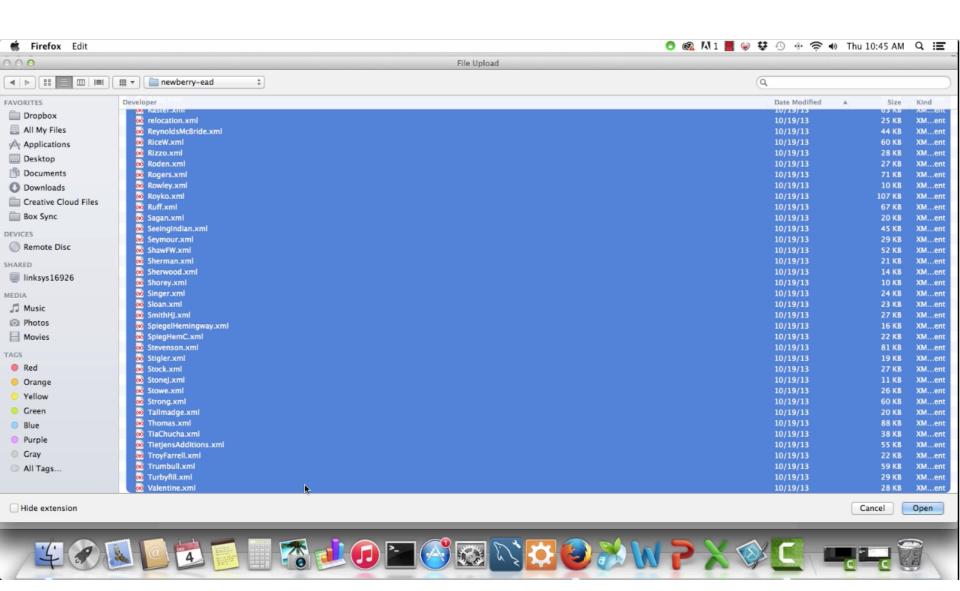


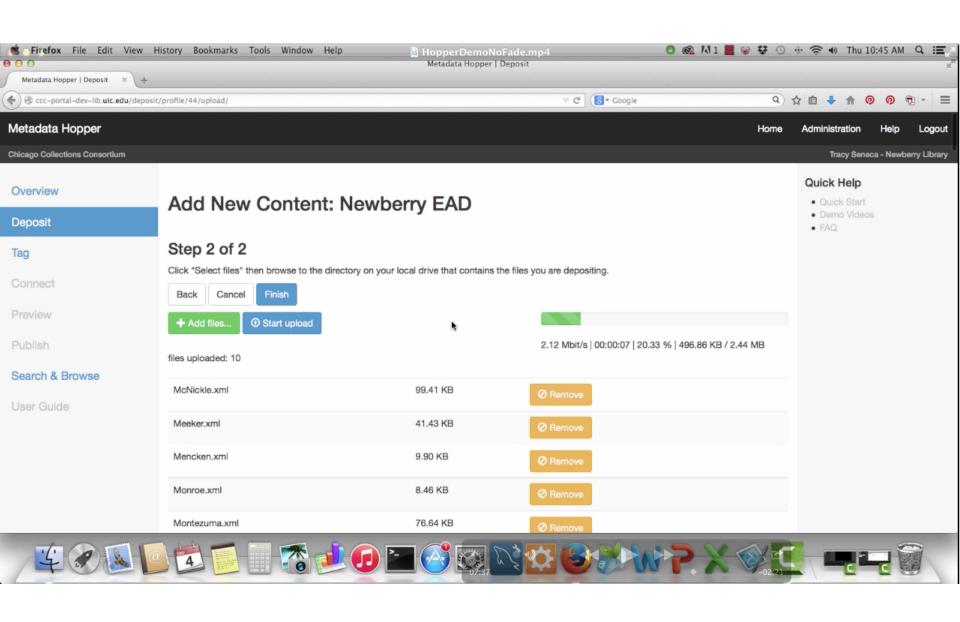


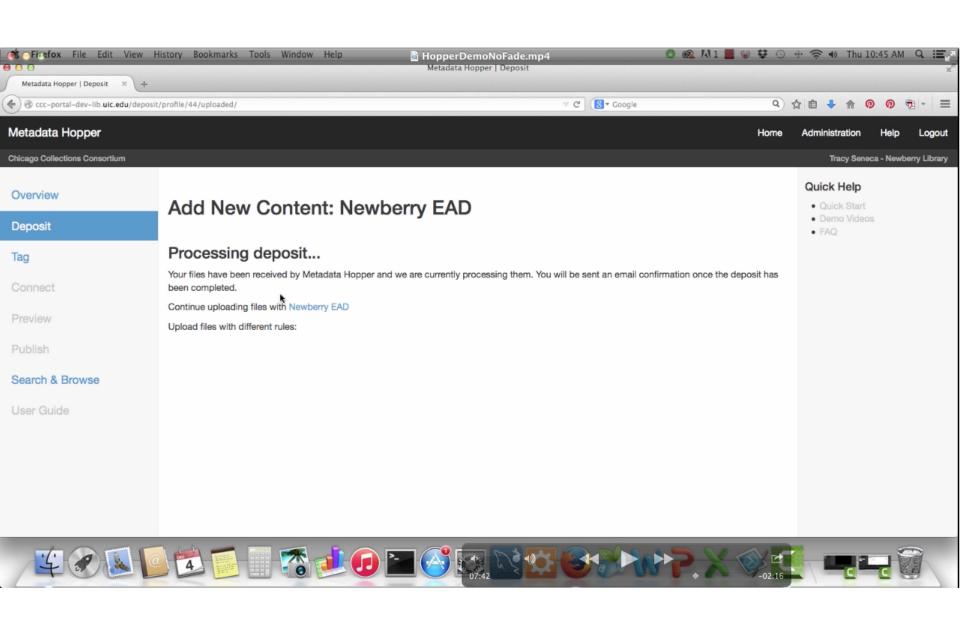












### Current Folder: INBOX















UIC WebMail FAQ

### Search Results | Unread | Delete

Forward | Forward as Attachment | Reply | Reply All

Subject: Chicago Collections Consortium Deposit Completed (0 errors)

From: ccc.portal.notifications@gmail.com Date: Tue, December 2, 2014 10:17 pm To: tjseneca@uic.edu (more)

Priority: Normal

Options: View Full Header | View Printable Version | Download this as a file | View Message Details | Add to Address Book

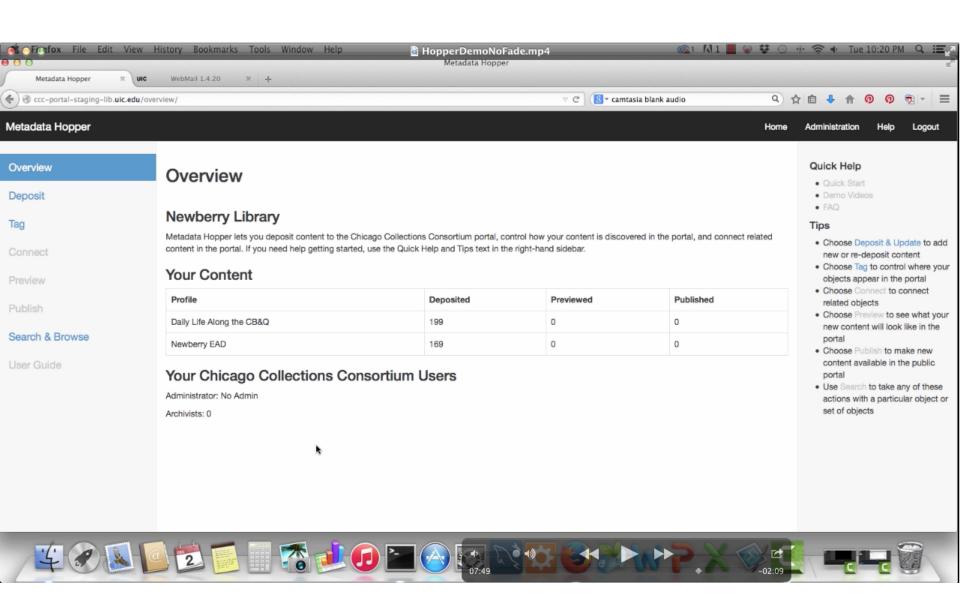
Deposit rules used: Newberry EAD Deposit requested by: Tracy Seneca Deposit requested: 12/03/2014 4:17 a.m.

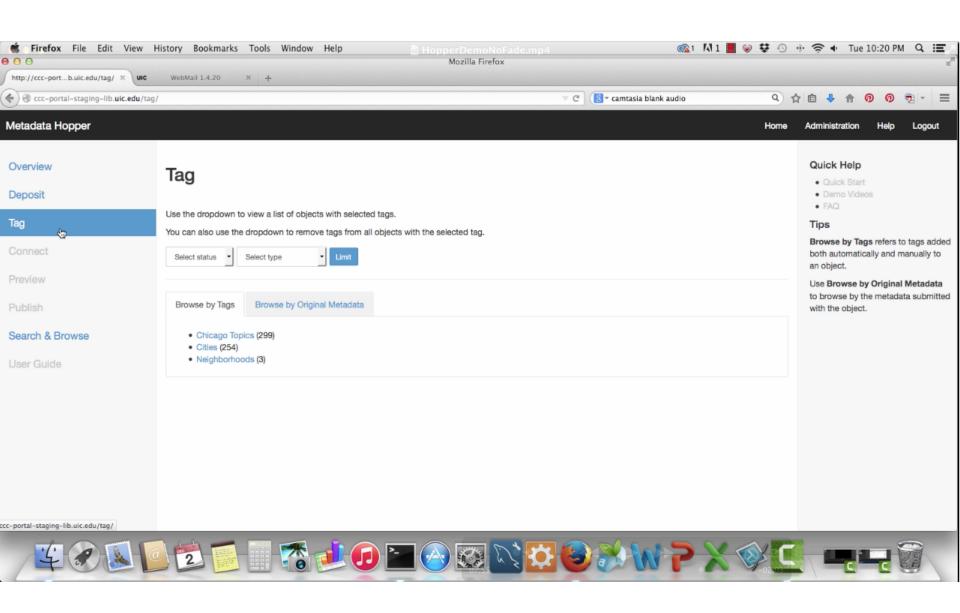
Deposited: 169 Deposit Failed: 0

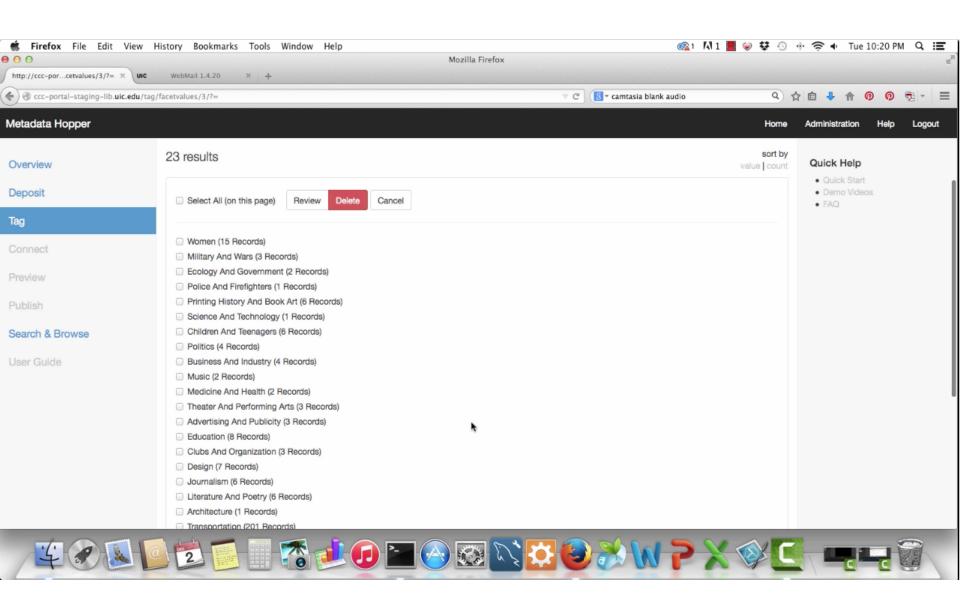
Log files have been created for any records that were not successfully uploaded. Contact the Portal Manager at for help if the error message in the log is not clear.

Errors may be due to unclosed XML tags, character encoding issues, or element tags and values using invalid characters, such as ':'. You will need to edit the XML file(s) and resubmit them.

### **Email Confirmation**

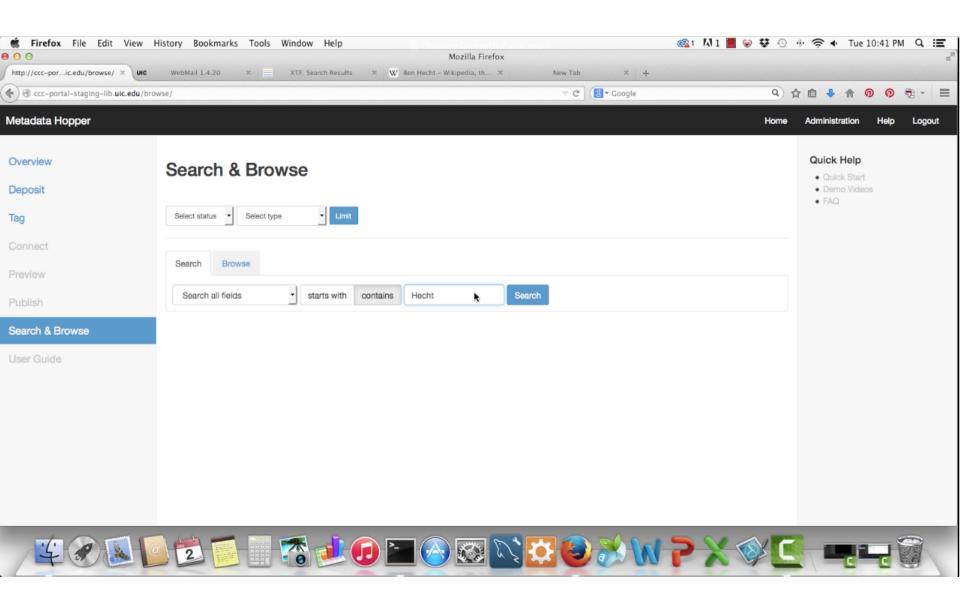


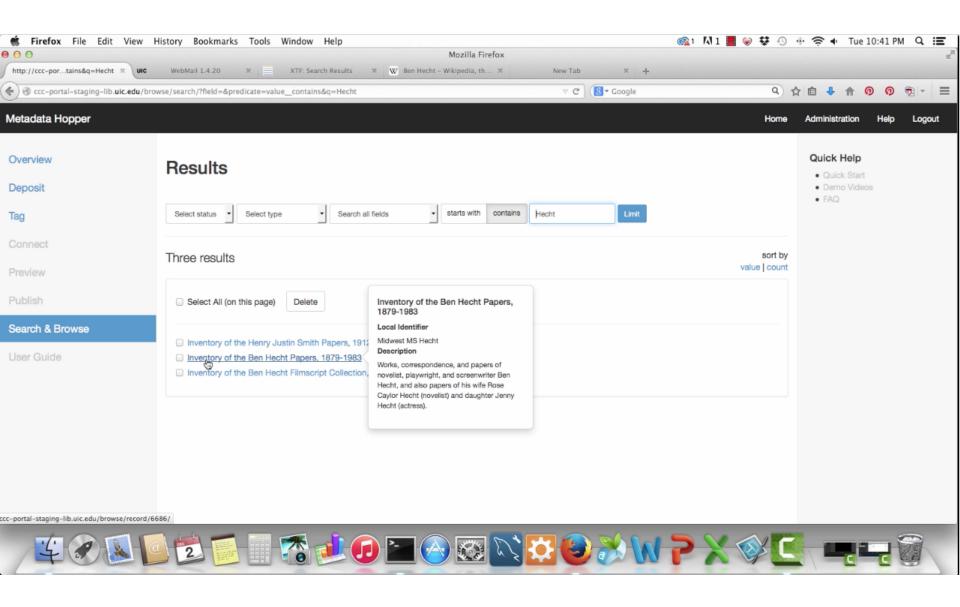


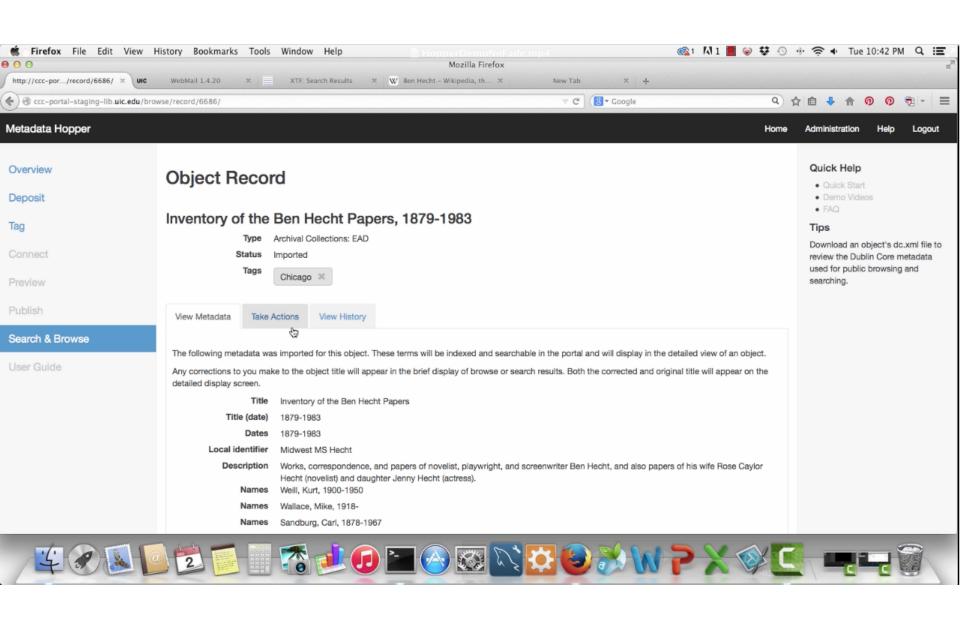


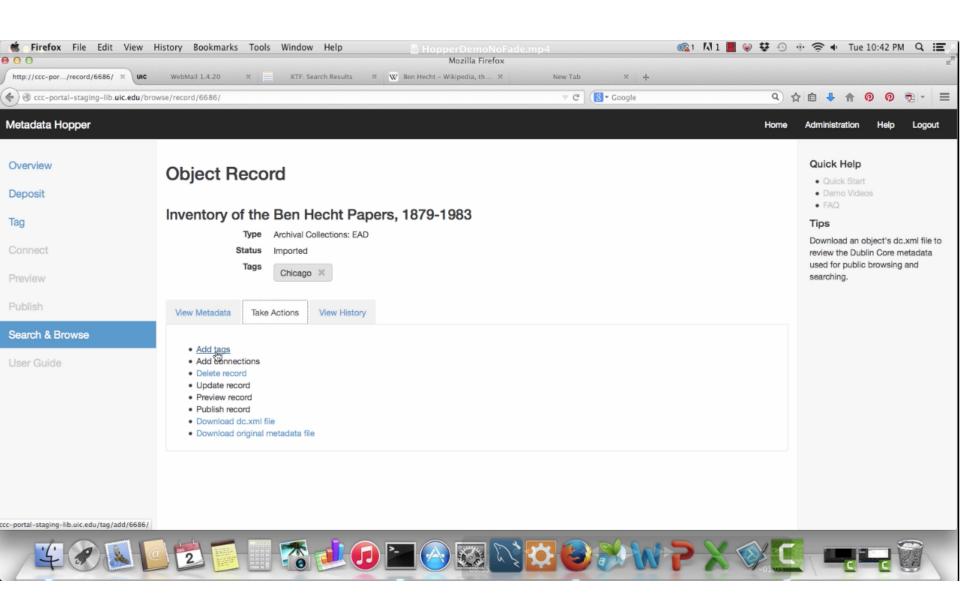
### Tagging: Individual Records

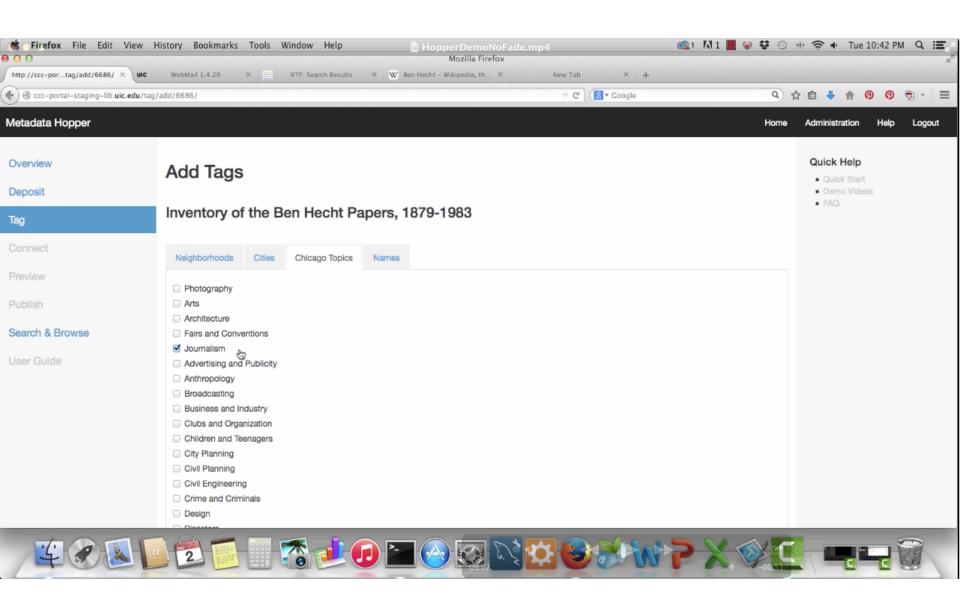
Newberry Library EAD Finding Aid

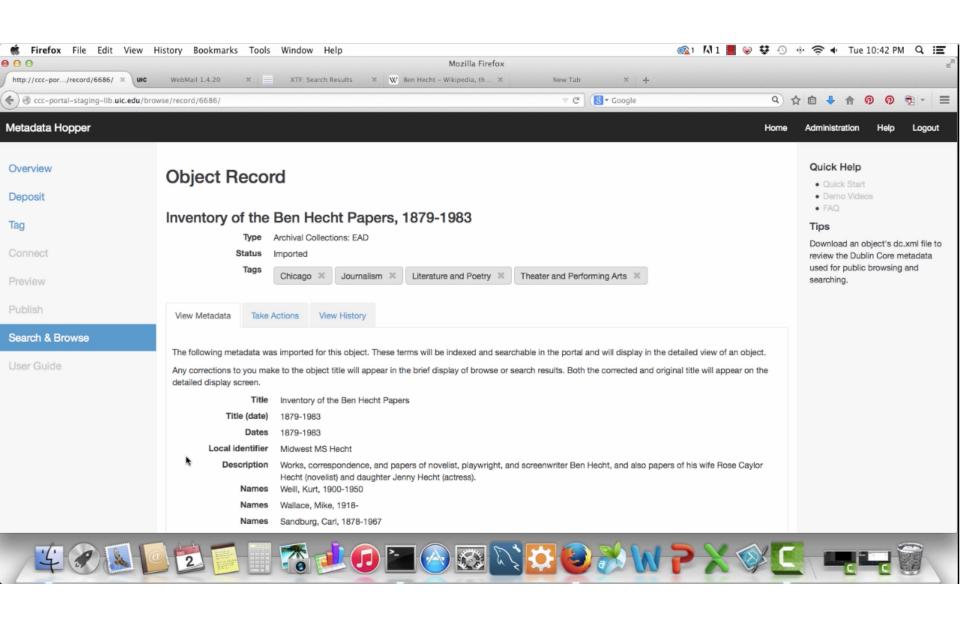






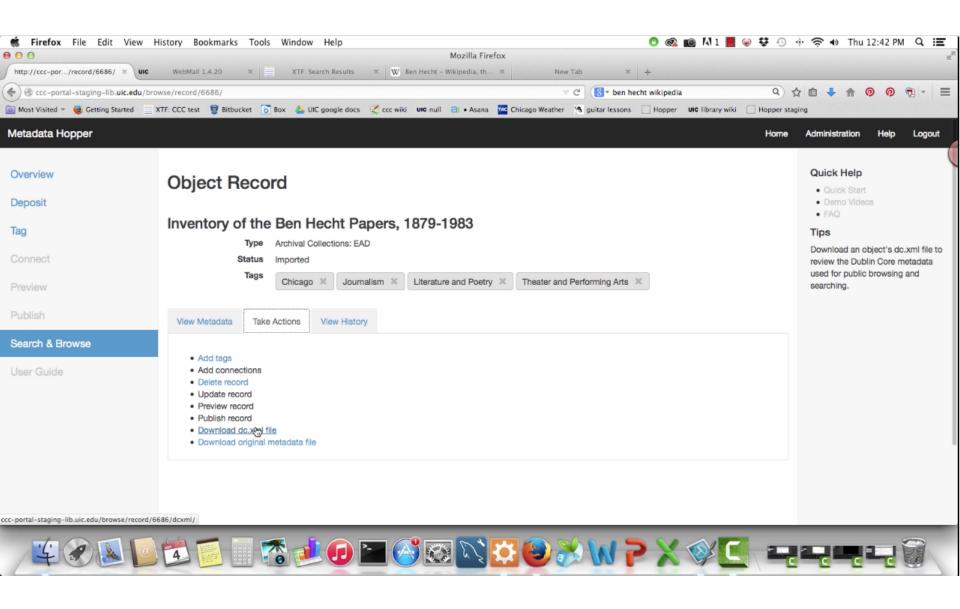


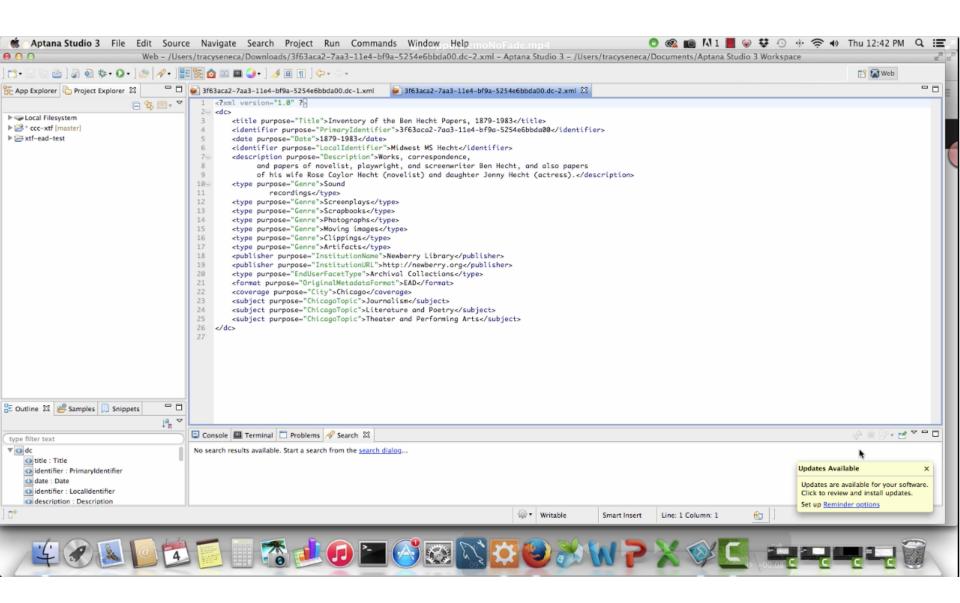




### The dc.xml File

Control of Front-End Browsing in XTF





### Important to note

- Open source
- Entirely configurable!
  - You define the mapping forms
  - You can set the tag lists
  - You create the autotagging rules
  - You can customize the dc.xml file output

### The Result?

### **Metadata Hopper Output**

(265)

(205)

(166)

(140)

(133)

(101)

(68)

(55)

(44)

Browse by: All Results: 1923 Items Sorted by: relevance Topic Architecture Transportation Women **Politics**  Ecology Clubs and Organization Ethnic and Cultural Groups (92) Design Arts · Children and Teenagers more Neighborhood

> (142) Loop Near West Side (21) Uptown (21)Near North Side (15)Kenwood (6) Near South Side (4) Woodlawn (4)Bridgeport (1) Lincoln Park (1) Pullman (1)more

### Type

 Archival Collections (1042) Digital Images

### Library/Archive

 University of Illinois at Chicago (1115) Newberry Library (368)(331)University of Chicago

 Art Institute of Chicago (108)Illinois Institute of Technology (1)

### Date

± 1950 (533) 1952 (531)

1951 (530)

± 1941 (528) 1945 (528)

more

Title: World's Columbian Exposition Photographs by C.D. Arnold, 1891-1894

Description: Nearly eight hundred large-format platinum photographic prints and twenty-five small-format

sepia photographic prints by C.D. Arnold, the official photographer of the 1893 World's Columbian Exposition in Chicago, illustrate the development of fair's site and building

Modify Search | New Search

Page: 1 2 3 4 5 ... Next

Browse by Facet | Title | Author

construction, the fair while ope...

Date: 1894

Title: Continental and Commercial National Bank Photographs Collection, c.1914

Description: Black and white photographs of the Continental and Commercial National Bank in Chicago,

Illinois, designed by D.H. Burnham and Company in 1912.

Date: 1914

Title: Bevis, Mildred B., Notebook of Designs for the Kalo Shop, c.1910-1913

Description: The notebook contains Arts and Crafts-influenced jewelry and tableware designs by Mildred Belle

Bevis Hanks for the Kalo Shop in Chicago, Illinois.

Date: 1913

Title: Adler, Dankmar, (1844-1900) Collection, 1844-1941

Description: Papers and photographs related to the life and work of Chicago structural engineer Dankmar

Adler, partner of architect Louis H. Sullivan.

Date: 1941 Subjects: Architecture

Title: Khan, Fazlur, (1929-1982) Collection, 1944-1988

Description: This collection holds professional papers and writings and a small group of personal papers of

the Bangladeshi structural engineer Fazlur R. Khan.

Date: 1988 Subjects: Design

Title: Weber, Peter J., (1863-1923) Papers, 1888-c.1945

Description: Photographs, correspondence, and printed matter documenting a selection of German-American

architect Peter J. Weber's built and unbuilt projects in the United States and abroad, though predominantly in the Chicago area. Projects represented in this collection span from his early

years as a student a ...

Date: 1922 Subjects: Architecture

Title: Sailor, Homer Grant, (1887-1968) Papers, 1914-1993

Description: Photographs and papers relating to the work of Chicago Prairie School architect Homer Grant

Sailor. Date: 1993

Subjects: Architecture | Design

### **Metadata Hopper Output**

Search: Transportation in subject [X]

Digital Images in type [X]

Results: 199 Items

Sorted by: relevance Go! Modify Search | New Search

Browse by Facet | Title | Author

Page: 1 2 3 4 5 ... Next

### Topic

### less

Children and Teenagers (2)

Families

(1)IXI

Transportation Women

(3)

### Type

Digital Images [X]

### Library/Archive

Newberry Library (199)

### Date

± 1948 (199)

① (199)

1905 (198)

± 1906 (1)



Title: Armature repair department of Diesel shop

Description: Three workmen repairing armatures for diesel engines

Date: 1905

Subjects: Transportation



Title: Union Station-view of concourse showing information desk

Description: Women and men clustered around the information desk. In the foreground are two

servicemen, bags at their feet, dealing with papers in their hands. List of arriving trains

shown above the information desk.

Date: 1905

Subjects: Transportation



Title: Filling journal boxes with oil at car repair tracks

Description: Man filling journal boxes with oil at car repair tracks

Date: 1905

Subjects: Transportation



Title: Commuters waiting for suburban train

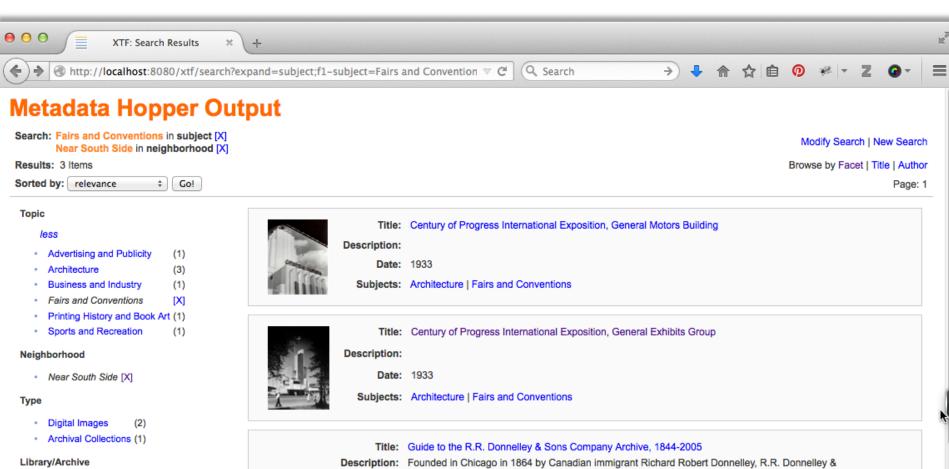
Description: Man folding a newspaper while standing on a train platform. Other people are visible on the

platform. There is a train in the distance under a signal tower (?), apparently coming into

the station. Parked cars are visible near the station itself.

Date: 1905

Subjects: Transportation



Art Institute of Chicago (2)

University of Chicago (1)

### Date

± 1932 (3)

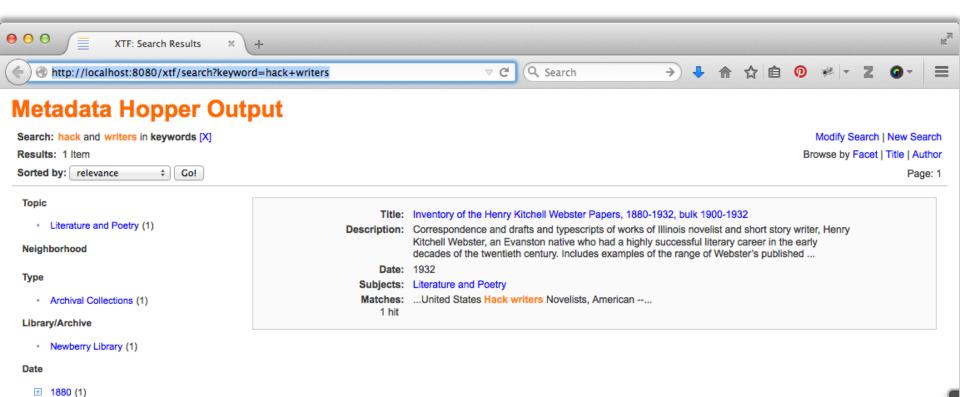
Sons Company is a leading global provider of printing and print-related services. The archive contains business records, product samples and promotional material, biographical files and

personal papers, histo...

Date: 2005

Subjects: Architecture | Fairs and Conventions | Advertising and Publicity | Business and Industry | Printing

History and Book Art | Sports and Recreation



Comments? Questions? | Choose Language

 ⊕ 1881 (1)
 ⊕ 1882 (1)
 ⊕ 1883 (1)
 ⊕ 1884 (1)
 more

## Thank you!! And thanks to:

- Andrew W. Mellon Foundation
- Kate Flynn, CCC Portal Manager
- Esther Verreau, Lead Developer
- Allan Berry, CCC Developer
- CCC Portal Committee Members
- Jaclyn Grahl, CCC Executive Director

chicagocollectionsconsortium.org