Pilot:

Cambridge Historical Commission Building Files

Digitize > Catalog > Expose > Explore > Preserve

pbc GIS cultivating spatial intelligence®

December 18, 2017

Cambridge Building Files Collection:
Digitize, Catalog, Expose, Explore, and Preserve

The Cambridge Historical Commission conducted a building survey in the early 1970s. A photo was taken and a survey form completed for each of the 16,000 buildings in the city. Each form was placed in a manila folder in a file cabinet. Over the years additional information was placed into these folders as additional research was conducted on particular buildings as part of the commission's official business or as other information turned up in newspaper articles or a variety of other documents which have been contributed or gathered by CHC staff. By now the Building Files collection occupies 16 four-drawer file cabinets. Each manilla folder is titled with a range of addresses.

This project, initiated by Charles Sullivan, Director of the Cambridge Historical Commission, and Emily Gonzalez, CHC Archivist, is intended to

- Develop a workflow for scanning and cataloging the documents
- Create a searchable catalog for the items
- Work with Cambridge GIS to incorporate building files which will allow the documents to be linked with other documents about buildings and for exploring the documents on historic maps.

In addition to accomplishing these practical capabilities, the project is oriented by several concerns related to long-term information stewardship:

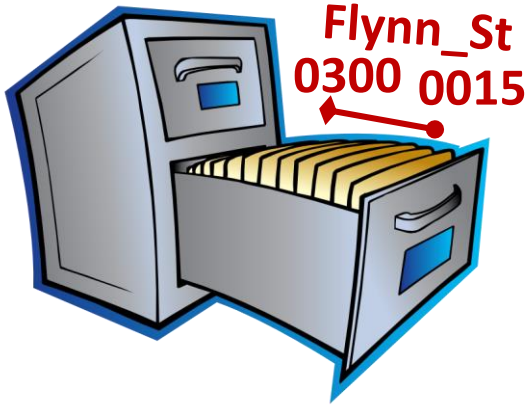
Assure that digital resources and associated documentation are secure against mishaps, natural disasters and natural decay.
Assure that resources and associated metadata are always interoperable between a variety of standards-based repository tools.
Provide a means of entering and preserving new information, including artifacts that are born digital.
Look forward to providing a platform for linking information form the Building Files collection to web-based applications and scholarship, including applications such as the Massachusetts Cultural Resource Information System.

Participants:

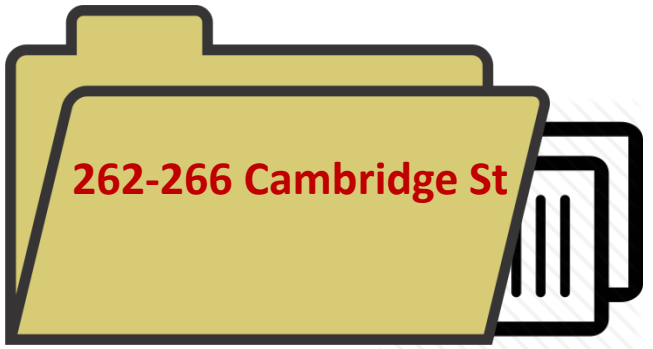
Charles Sullivan, Director, Cambridge Historical Commission
Emily Gonzalez, Cambridge Historical Commission Archivist
Meta Partenheimer, Sarah Burks, Cambridge Historical Commission
Jeff Amero, and Katherine Grillo Cambridge GIS
Peter Stott, Massachusetts Cultrual Resouces Information System
Paul Cote, pbcGIS.com

Folders and Items (Existing)

64 File drawers of documentation about Cambridge Buildings



16,000 Folders, each more or less related to a property or building



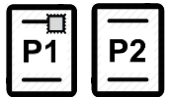
Items: 000, 001

1. Put Items and Pages in Order

Prepare batches of folders for scanning.

Cover Sheet for Each folder

Multi-Page Documents, Photo Front and Reverse



Get advice from vendors on pre-organization.

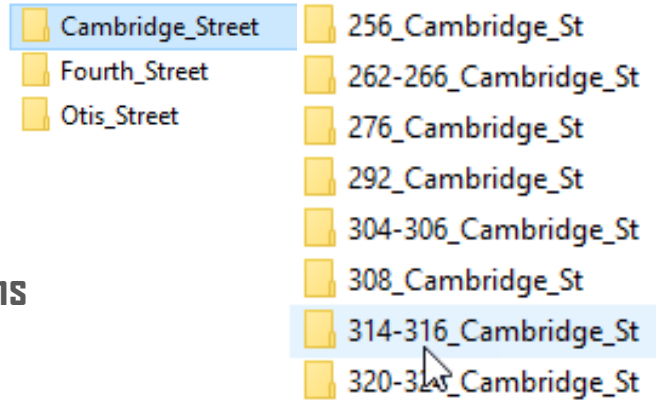
Scanning Workflow & Naming Scheme (ideas)

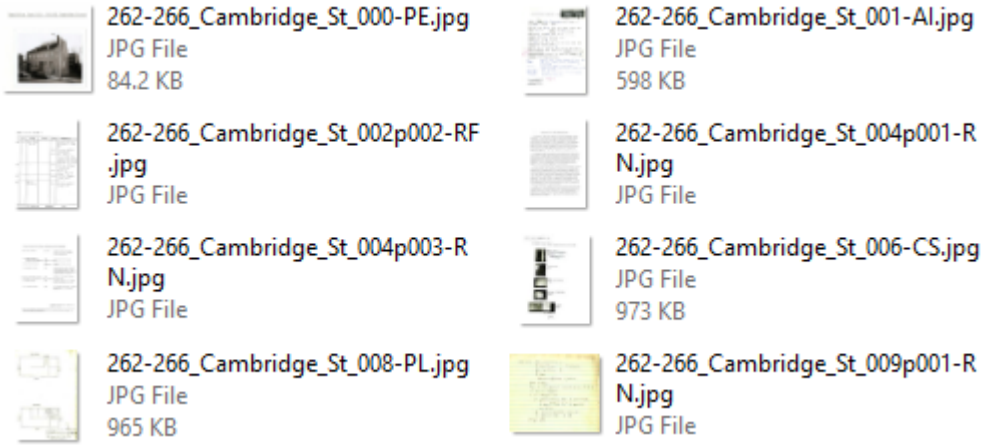
Rough Sort of Item Types:
Categories could be applied by quick visual inspection during, or after? scanning process.

- AI: Architectural Inventory Form
- RF: Research Form
- PE: Exterior View
- PI: Interior Detail
- AM: Article or Monograph
- RN: Research Notes
- OB: Official Business
- CS: Contact Sheet
- PL: Plan
- LT: Letter
- UC: Unclassified
- HS: Historic Building Survey

Rough Sort of Item Types:
May also be accomplished in a QA process as batches of scans are checked and uploaded to Omeka

ScanID = 264-266_Cambridge_St_001p001-AI.tif





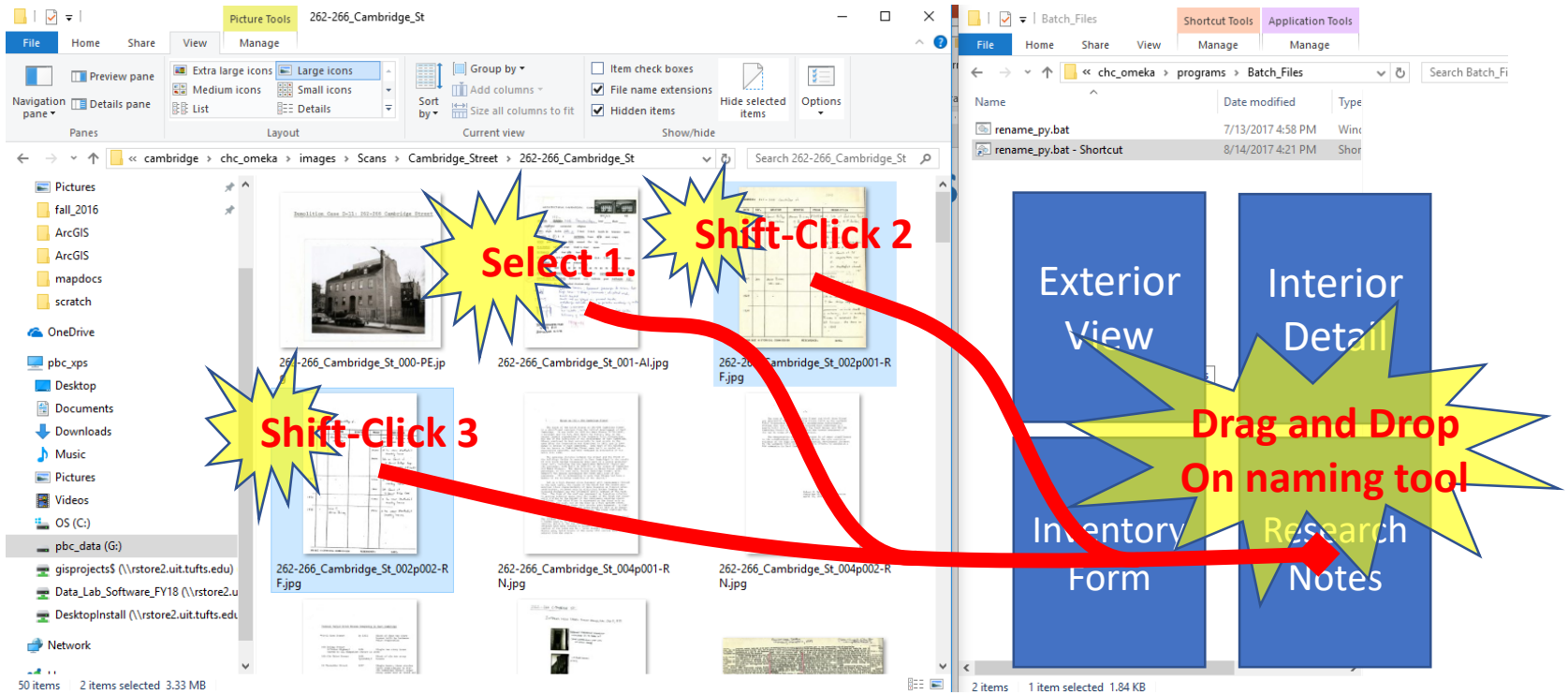
Tools for Post-Scanning Rough Sort

Drag and Drop Item Type Assignment.

Open Folder with Large Tiles View

Visually select files of a particular type.

Select many, then drag and drop onto the appropriate Type Changer tool



Geocode Addresses and Supplement Master Address List

(Using & Extending Cambridge Master Address file)

A script creates a table of all addresses suitable for geocoding in ArcMap

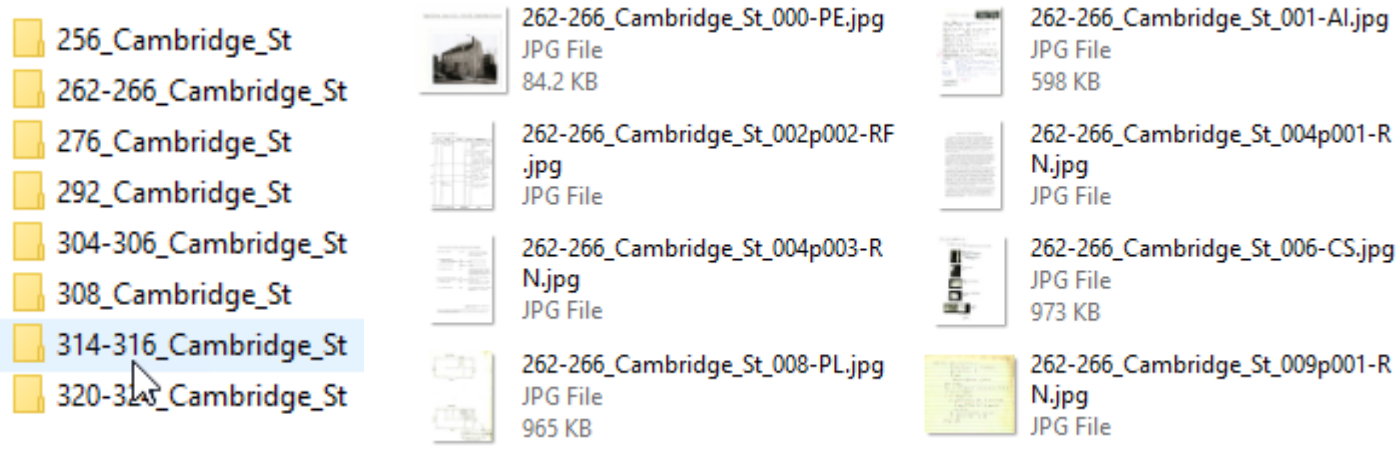
ArcMap Geocoding Process is flexible regarding spelling.

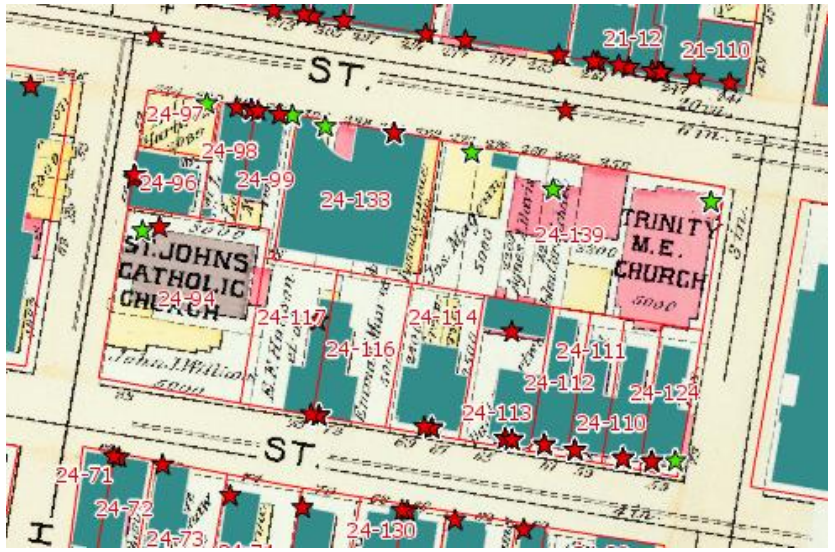
Addresses not matched are added to table of Historical address points.

Future geocoding is made using a union of the official address points and the CHC Historical Address Points

Result is a dictionary of all CHC Addresses with:

- Latitude
- Longitude,
- Map-Lot
- Building_ID





Shape	Full_Addr	BldgID	ml	POINT_X	POINT_Y
Point	39 Fourth St	416-31	23-36	-71.081493	42.371644
Point	113-A Fourth St	530-10	26-81	-71.08247	42.368326
Point	111 Fourth St	530-6	26-80	-71.082266	42.368371
Point	113 Fourth St	530-12	26-81	-71.082312	42.368311
Point	119-1/2 Fourth St	530-31	26-151	-71.08254	42.368102
Point	117 Fourth St	530-24	26-150	-71.082433	42.368817
Point	105 Fourth St	507-34	26-59	-71.082241	42.368683
Point	106 Fourth St	514-19	24-2	-71.081914	42.36875
Point	104 Fourth St	514-16	24-3	-71.081913	42.368833
Point	118-R Fourth St	533-14	17-22	-71.081987	42.368169
Point	106-R Fourth St	514-19	24-2	-71.081821	42.368741
Point	23 Fourth St		22-142	-71.081455	42.372479

Data Round-Up Scripts

For each image file in the scans folder:

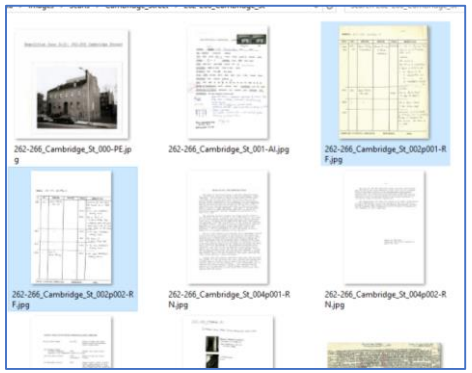
Use the item-type dictionary to apply meaningful titles and keywords

Use the Address Points dictionary to look up the coordinates, building-ID and Map-Lot.

Write a comma-delimted table formatted for the Omeka CSV import tool

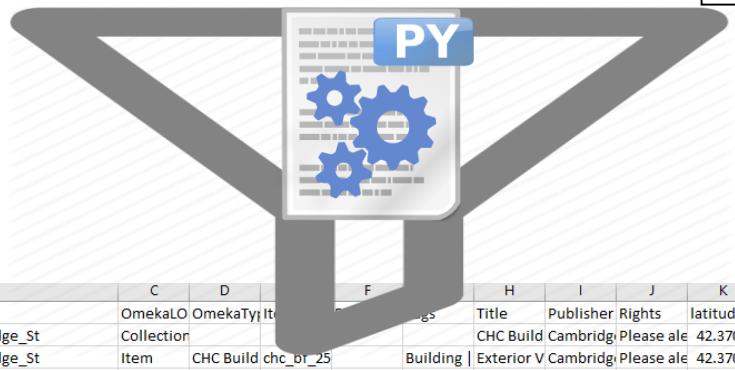
Future version could:

- Embed the metadata and coordinates into the image files themselves.
- Import the collection into any content management system that permits systematic import.



Shape	Full_Addr	BldgID	ml	POINT_X	POINT_Y
Point	39 Fourth St	416-31	23-36	-71.081493	42.371644
Point	113-A Fourth St	530-10	26-81	-71.08247	42.368326
Point	111 Fourth St	530-6	26-80	-71.082266	42.368371
Point	113 Fourth St	530-12	26-81	-71.082312	42.368311
Point	119-1/2 Fourth St	530-31	26-151	-71.08254	42.368102
Point	117 Fourth St	530-24	26-150	-71.082433	42.368817
Point	105 Fourth St	507-34	26-59	-71.082241	42.368683
Point	106 Fourth St	514-19	24-2	-71.081914	42.36875
Point	104 Fourth St	514-16	24-3	-71.081913	42.368833
Point	118-R Fourth St	533-14	17-22	-71.081987	42.368169
Point	106-R Fourth St	514-19	24-2	-71.081821	42.368741
Point	23 Fourth St		22-142	-71.081455	42.372479

- AI: Architectural Inventory Form
- RF: Research Form
- PE: Exterior View
- PI: Interior Detail
- AM: Article or Monograph
- RN: Research Notes

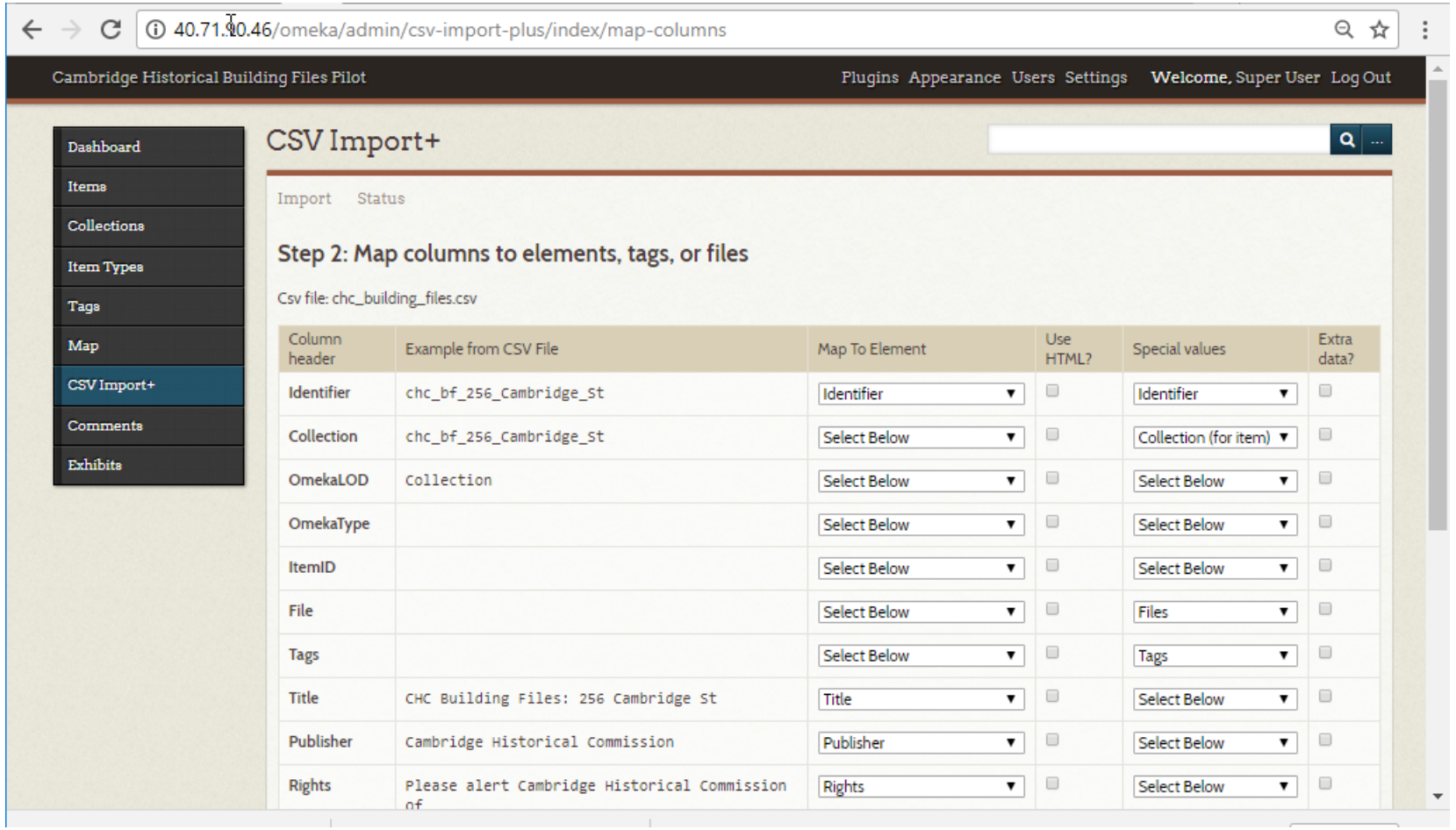


	A	B	C	D	F	H	I	J	K	L	M	N	O	P
1	Identifier	Collection	OmekaLO	OmekaTy		Title	Publisher	Rights	latitude	longitude	MapLot	BuildingID	Address	CHCDocT
2	chc_bf_256_Cambridge_St	chc_bf_256_Cambridge_St	Collection			CHC Build	Cambridg	Please ale	42.37092	-71.08	24-139		256 Cambi	Collec
3	chc_bf_256_Cambridge_St_000	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Building |	Exterior V	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Exterio
4	chc_bf_256_Cambridge_St_001	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Research	Research	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Researc
5	chc_bf_256_Cambridge_St_002	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
6	chc_bf_256_Cambridge_St_003	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Photograp	Contact Sh	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Contact S
7	chc_bf_256_Cambridge_St_004	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
8	chc_bf_256_Cambridge_St_005	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Research	Research	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Researc
9	chc_bf_256_Cambridge_St_006	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Research	Research	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Researc
10	chc_bf_256_Cambridge_St_007	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
11	chc_bf_256_Cambridge_St_008	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article o
12	chc_bf_256_Cambridge_St_009	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
13	chc_bf_256_Cambridge_St_010	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
14	chc_bf_256_Cambridge_St_011	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
15	chc_bf_256_Cambridge_St_012	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
16	chc_bf_256_Cambridge_St_013	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article o
17	chc_bf_256_Cambridge_St_014	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
18	chc_bf_256_Cambridge_St_015	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Research	Research	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Researc
19	chc_bf_256_Cambridge_St_016	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Article | N	Article or	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Article
20	chc_bf_256_Cambridge_St_017	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Architectu	Site Plan:	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Site Pl
21	chc_bf_256_Cambridge_St_018	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Building |	Exterior V	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Exterior
22	chc_bf_256_Cambridge_St_020	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Building |	Exterior V	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Exterio
23	chc_bf_256_Cambridge_St_021	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Building |	Exterior V	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Exterio
24	chc_bf_256_Cambridge_St_022	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Uncategor	Unccatego	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Unccate
25	chc_bf_256_Cambridge_St_023	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Historical	Letter: 25(	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Letter
26	chc_bf_256_Cambridge_St_024	chc_bf_256_Cambridge_St	Item	CHC Build	chc_bf_25	Historical	Letter: 25(	Cambridg	Please ale	42.37092	-71.08	24-139	256 Cambi	Letter

Import Images and Rough Metadata On-Line Repository (Omeka)

Images and CSV File are uploaded to a cloud server which also has Omeka Installed.

The Omeka Import CSV + Plugin is used to import the images and to create a simple catalog record for each..



Omeka Tour: What is Omeka?

Omeka is a free and open-source project of the Roy Rosenzweig Center for History and New Media, George Mason University.
www.Omeka.org

Huge community of users and contributors, coders and consultants

Engagement with serious library and cultural preservation projects like Digital Commonwealth and Public Library of America

Omeka as a Framework for Attaching Knowledge to Digital Resources

Visit:
http://40.71.90.46/omeka

Omeka data base is organized as a hierarchy:

Collection = Building Files Folder

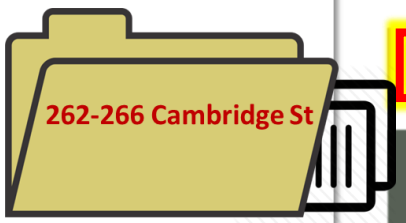
Items are folder contents, can be multi-paged.

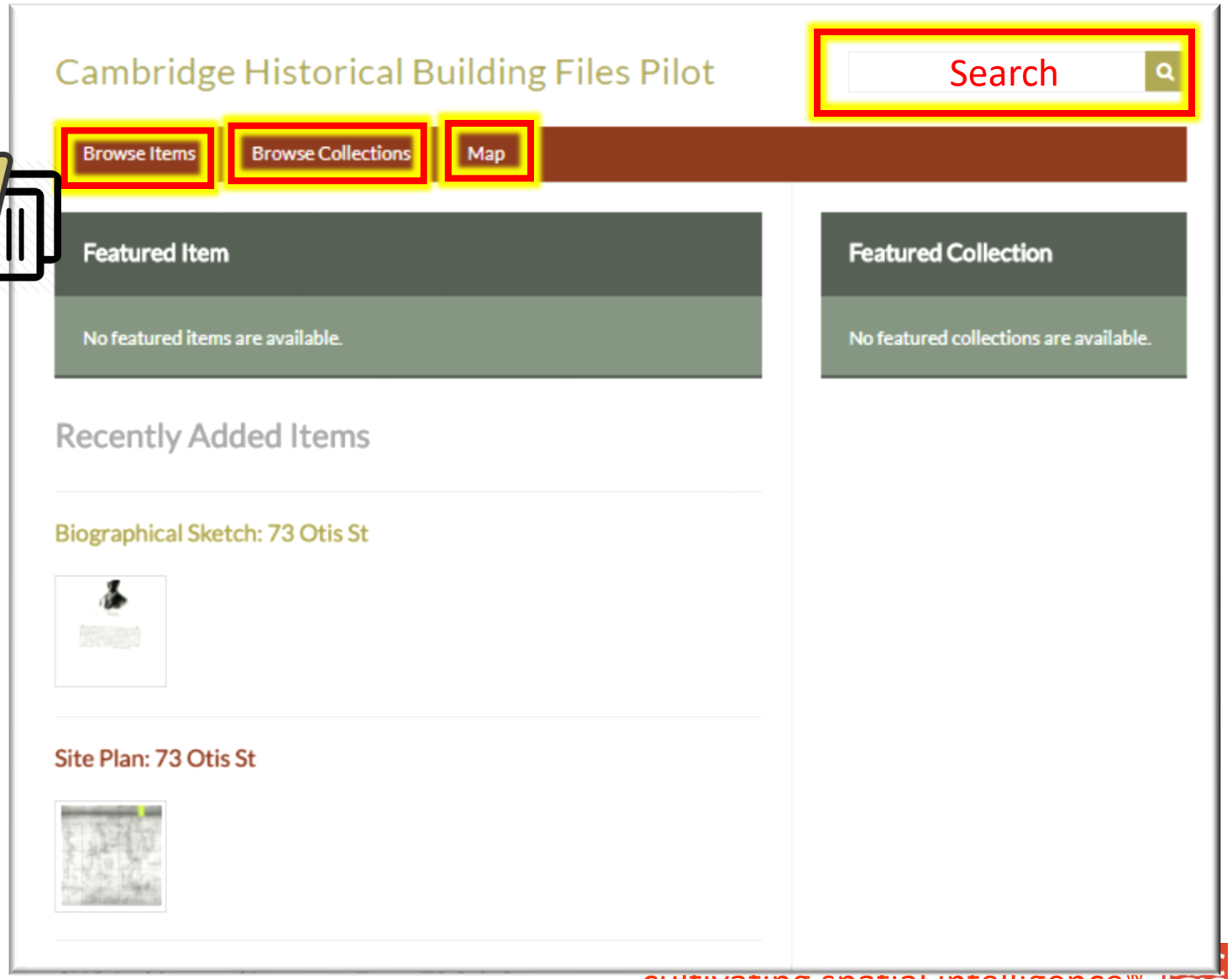
Try search by Map

Try searching by "Title Contains 'Form'"

Try searching by "Title Contains 'Form'"

Try searching by "Title Contains 'Form'"





Play with Public Omeka Site

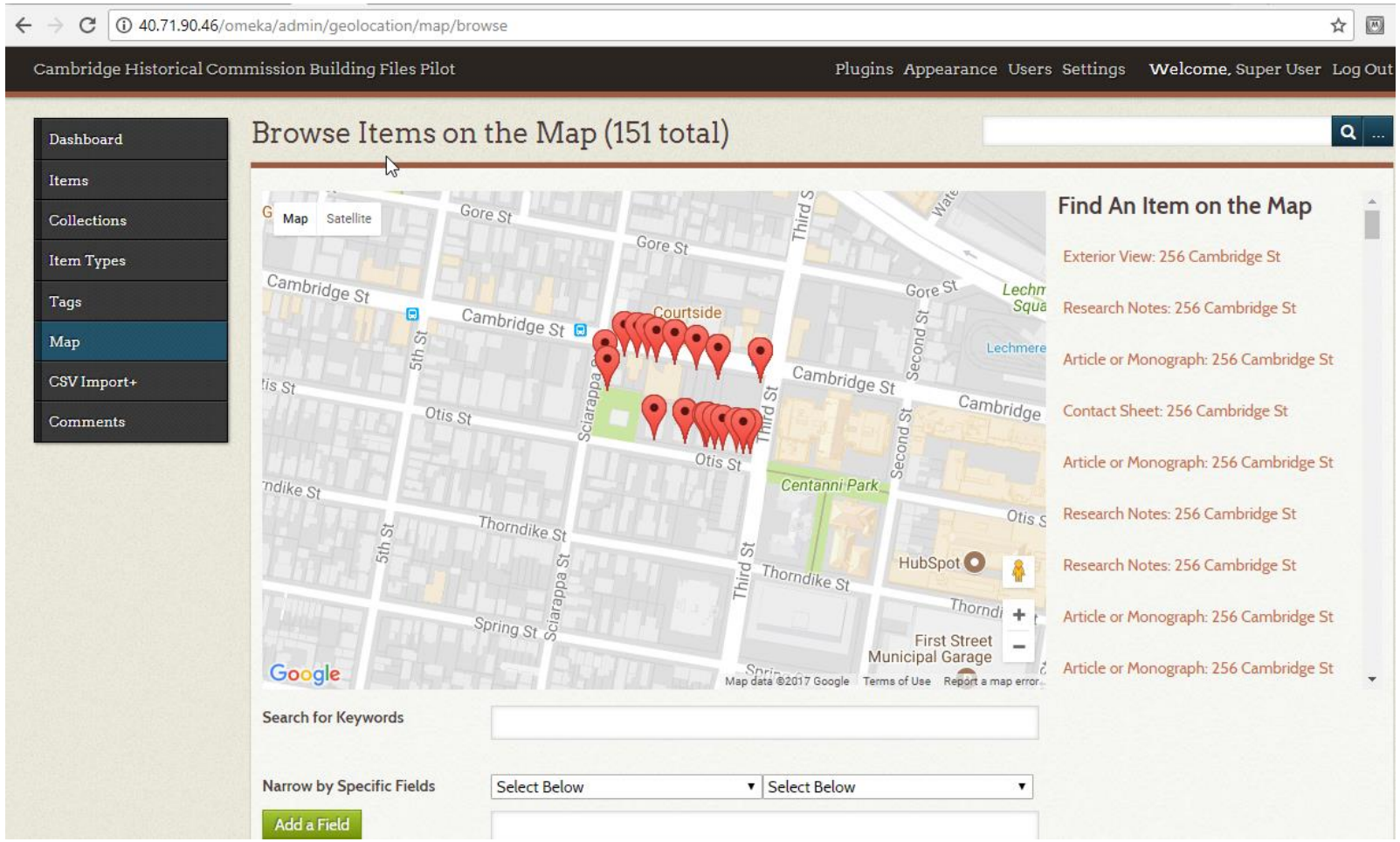
Visit:
http://40.71.90.46/omeka

Click "Map" on the left-hand sidebar to explore items on the map.

This Omeka site is bare-bones for the pilot, but can be made fancier with some free plug-ins see the Cleveland Historical site.

Different levels of access can be established for groups of users.

I'll create "Researcher" accounts for each of you. After some training, Admin role can be assigned.



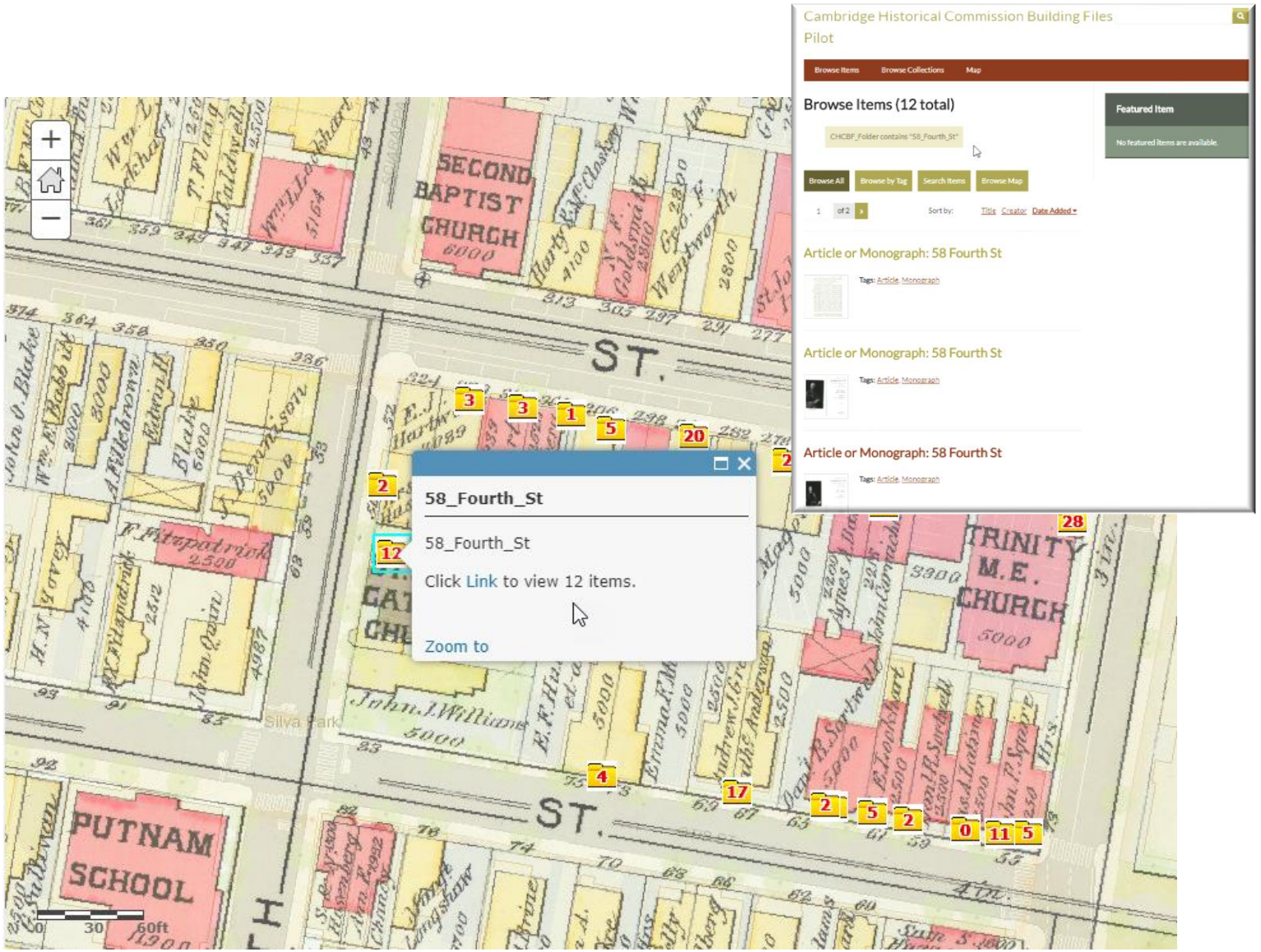
Play with Pilot Web GIS Application

Elements of the same CSV file used to create Items and Collections in the Omeka repository are used to put folders on a web-map from 1890.

Folders are displayed with a tag showing the number of items.

Click on a folder to reveal a URL to an Omeka page showing a thumbnail and title for each item from the slected folder.

This is a very basic GIS web-app that demonstrates how the Building Files items can be explored through GIS. There are many other possibilities.



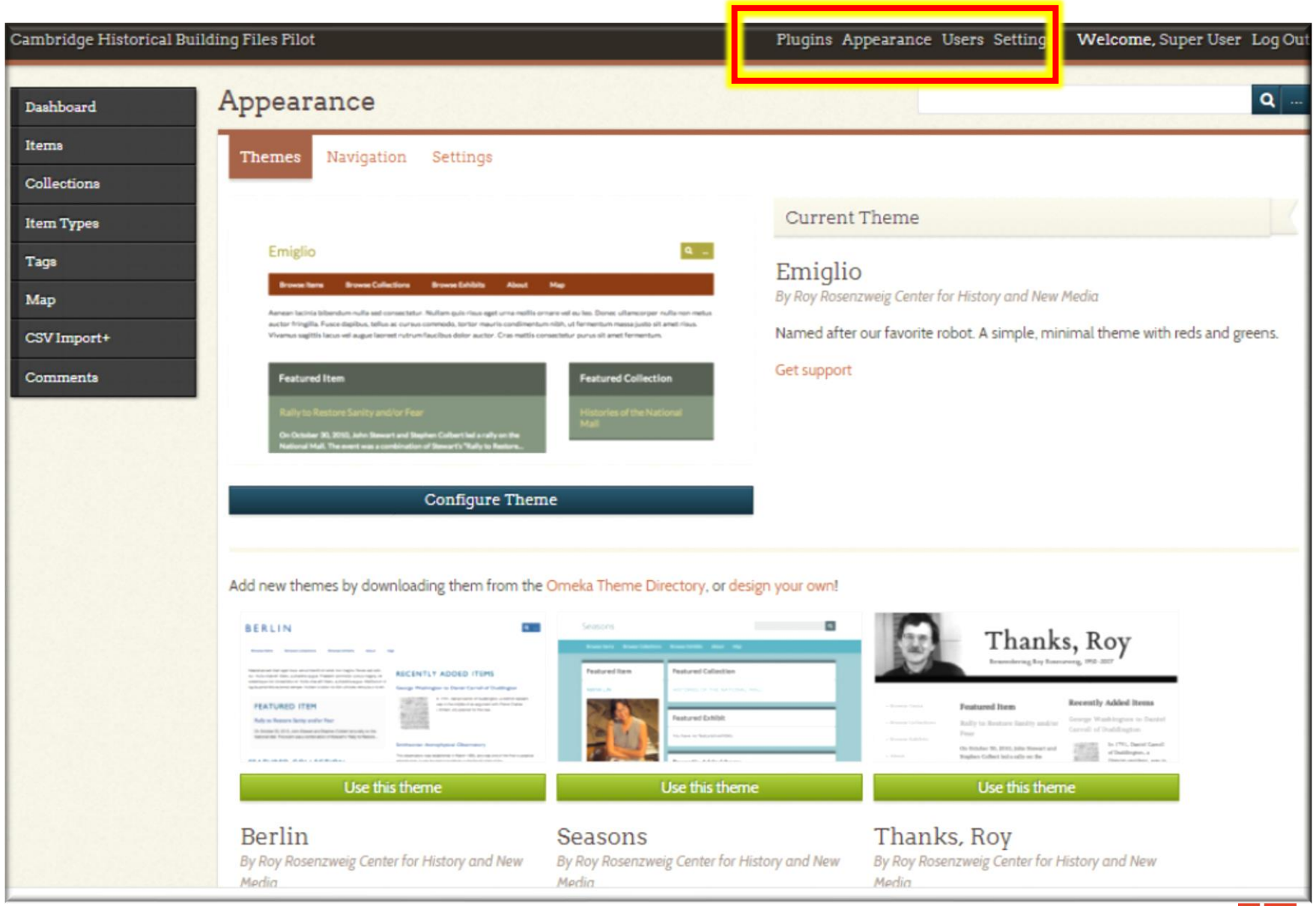
Omeka Tour: Super-User Tasks

Control look, feel and Branding of site using pre-fab "Themes"

Install Plug-Ins:

- Import CSV Plugin
- Library of Congress and Ghetty Vocabularies
- PDF Viewer / Export Plugin
- Neatline / Time Plugin
- CurateScape Plugin – make site like ClevelandHistorical.org

Add and Manage Users



Omeka Tour: Admin User Tasks

Create new Items

Upload Files (associate them with Items)

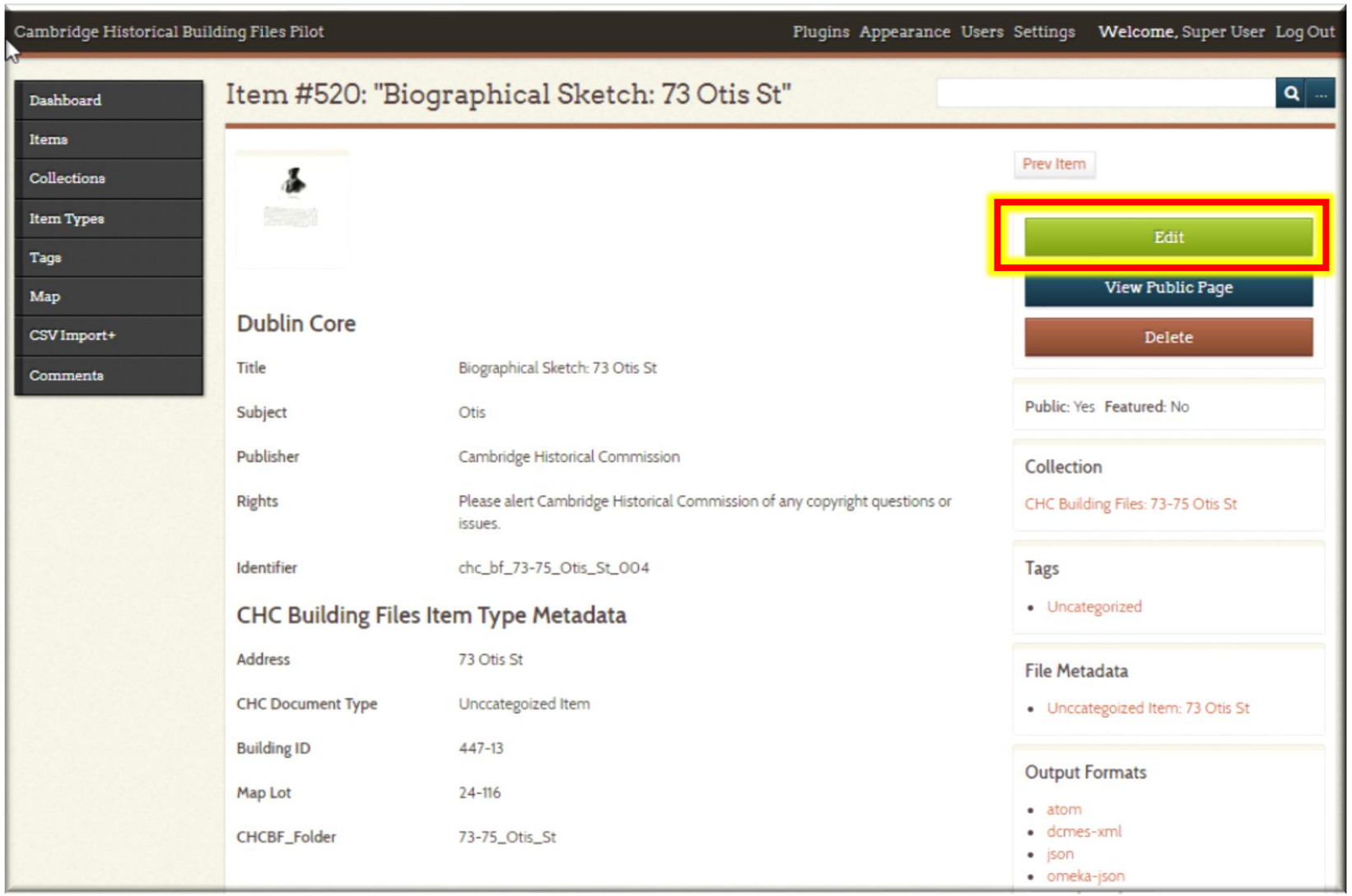
Edit Metadata and Location Info for Items

Create New Item Types

Create new metadata fields and controlled lists (MACRIS-ID)

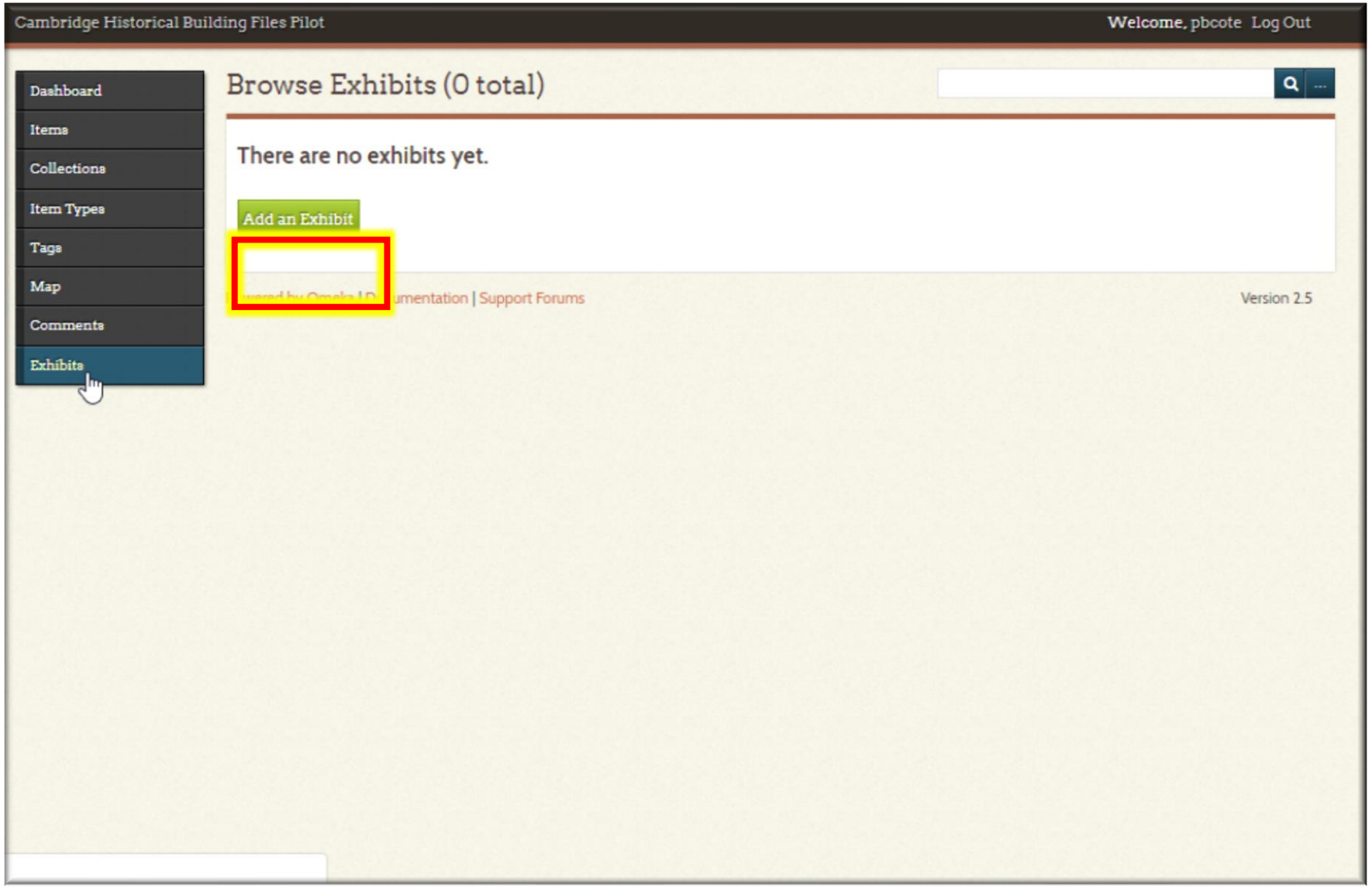
Inspect and Edit Tags Globally

More Omeka Documentation



Omeka Tour: Create exhibits that use items into ordered, annotated presentations.

Omeka Documentation for
[Exhibit Builder]



Omeka Tour: Export Archive / Backup CSV

After data has been enhanced in Omeka, Images and item metadata will need to be written out to files with an enhanced CSV catalog

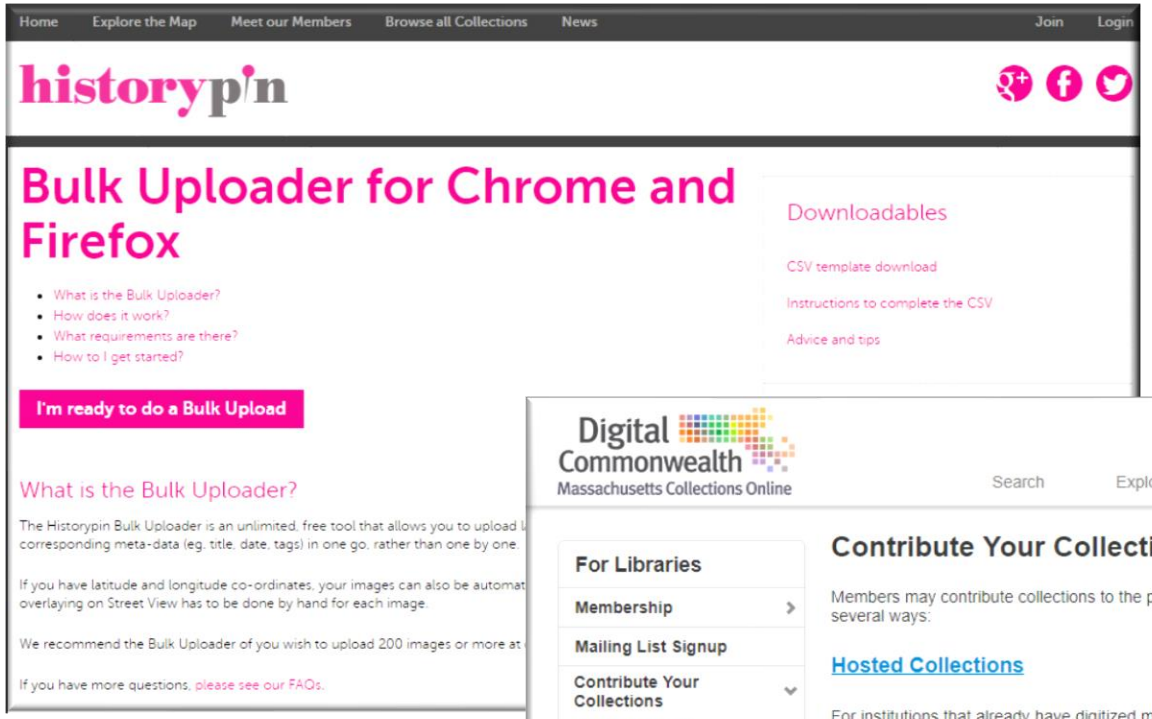
This wil be important for general back-up putposes

This will also be the pathway for importing into the Cambridge Document Managament System

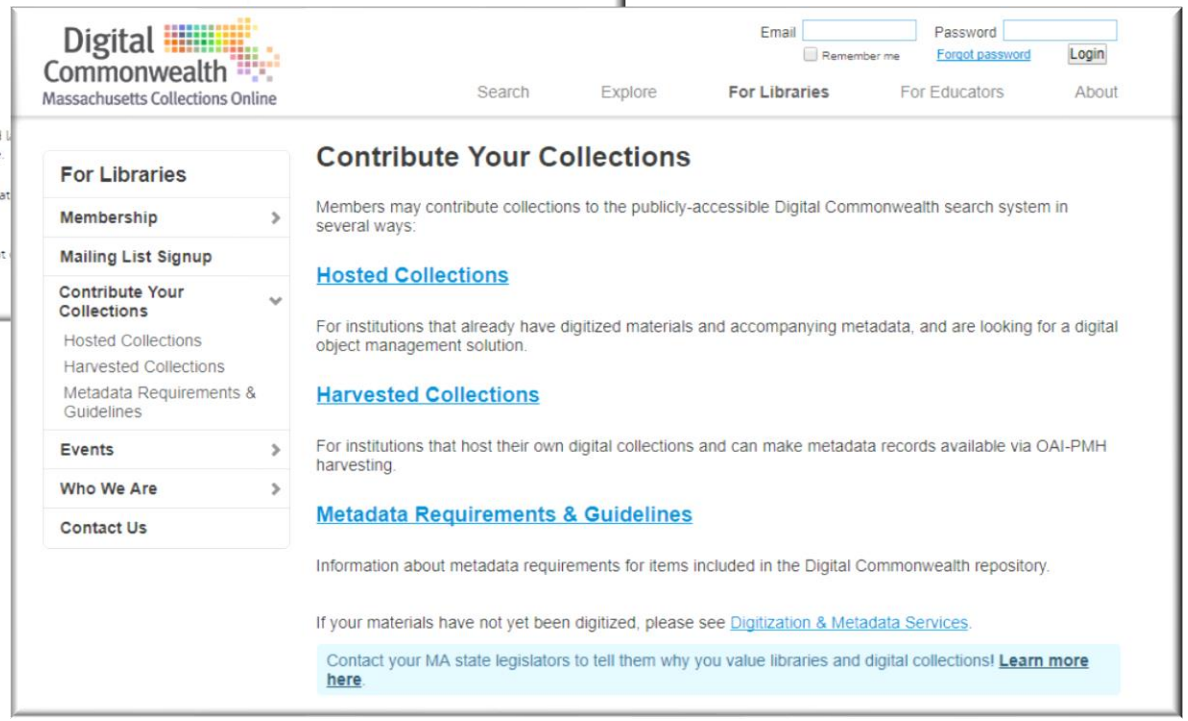
[Omeka Open Archive Initiative (OAI) Export Tool]

Omeka Tour: Export Selected Items and Metadata

[HistoryPIN Upload Tool]



Contribute Selected Items to [Digital Commonwealth]



Resources

These references provide some background on how to organize a digitizing project.

- Digital Public Library of America
  Digital Reformatting and File Management
  http://dp.la/info/about/projects/public-library-partnerships/digital-reformatting-and-file-management/
- National Archives:
  Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images
  https://www.archives.gov/files/preservation/technical/guidelines.pdf
- Smithsonian Institution
  Digitizing Collections
  https://siarchives.si.edu/what-we-do/digital-curation/digitizing-collections