

Pilot:

# Cambridge Historical Commission

## Building Files

Digitize > Catalog > Expose > Explore > Preserve



## **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:

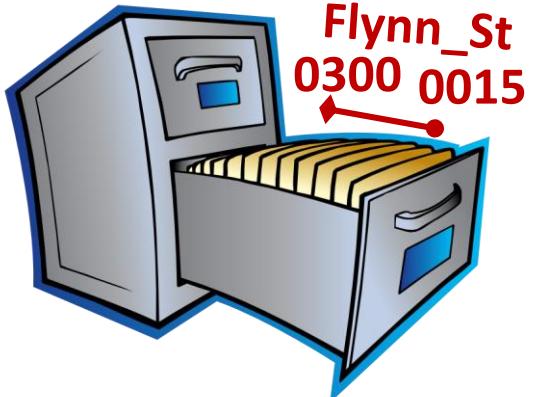
Assure that digital resources and associated documentation are secure against mishaps, natural disasters and natural decay.  
Assure that resources and associated metadata are always interoperable between a variety of standards-based repository tools.  
Provide a means of entering and preserving new information, including artifacts that are born digital.  
Look forward to providing a platform for linking information from 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 Cultural Resources 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.



## Rough Sort of Item Types:

May also be accomplished in a QA process as batches of scans are checked and uploaded to Omeka

[www.pbcGIS.com](http://www.pbcGIS.com)

- 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

**ScanID = 264-266\_Cambridge\_St\_001p001-AI.tif**

Cambridge\_Street  
Fourth\_Street  
Otis\_Street

256\_Cambridge\_St  
262-266\_Cambridge\_St  
276\_Cambridge\_St  
292\_Cambridge\_St  
304-306\_Cambridge\_St  
308\_Cambridge\_St  
314-316\_Cambridge\_St  
320-325\_Cambridge\_St

262-266_Cambridge_St_000-PE.jpg JPG File 84.2 KB	262-266_Cambridge_St_001-AI.jpg JPG File 598 KB
262-266_Cambridge_St_002p002-RF.jpg JPG File	262-266_Cambridge_St_004p001-RN.jpg JPG File
262-266_Cambridge_St_004p003-RN.jpg JPG File	262-266_Cambridge_St_006-CS.jpg JPG File 973 KB
262-266_Cambridge_St_008-PL.jpg JPG File 965 KB	262-266_Cambridge_St_009p001-RN.jpg JPG File

cultivating spatial intelligence®

pbc  
GIS

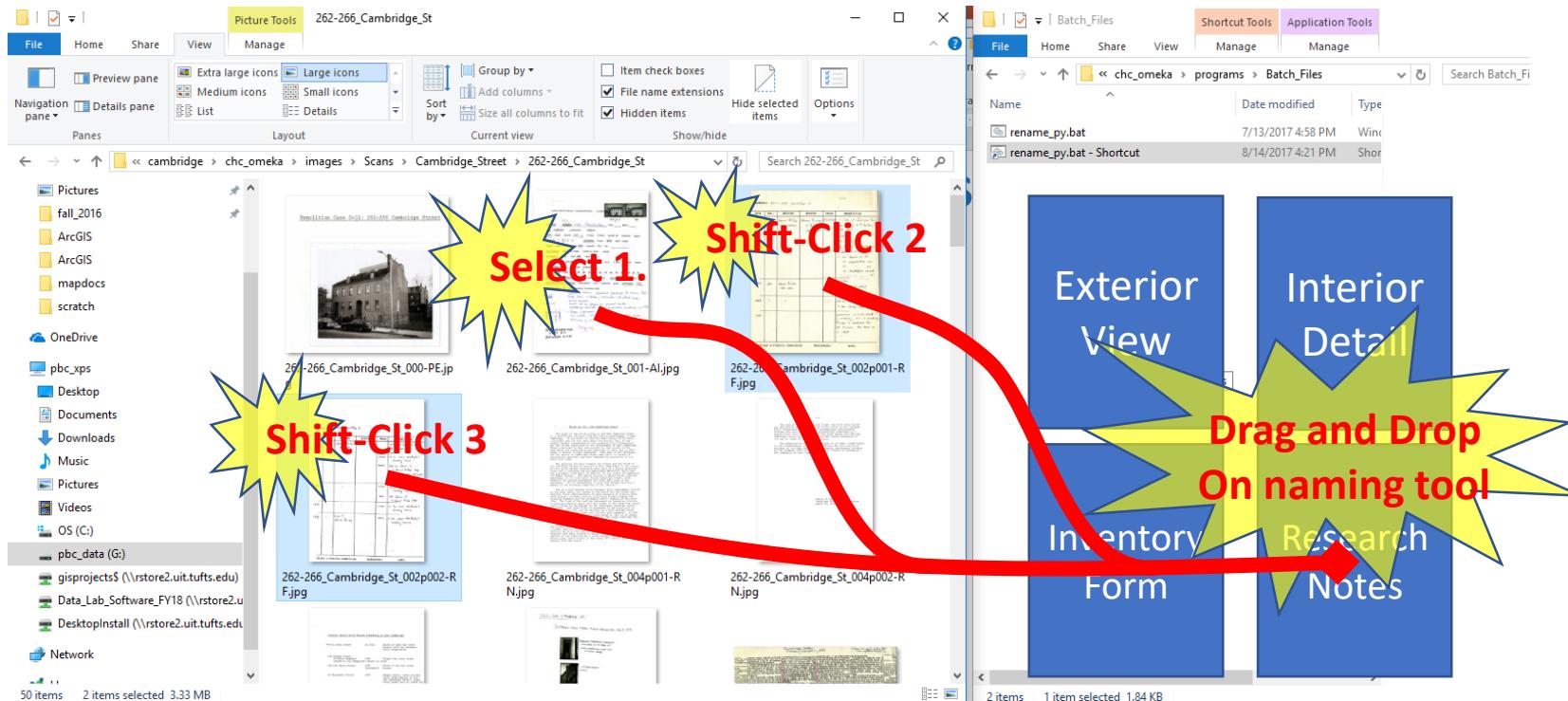
# Tools for Post-Scanning Rough Sort

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

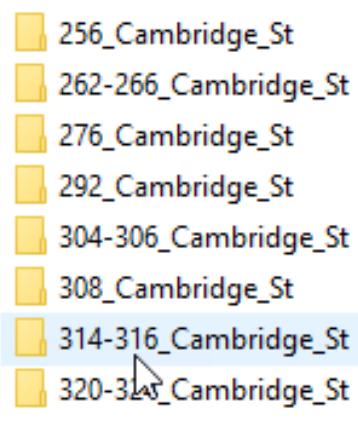
ArcMap Geocoding Process is flexible regarding spelling.

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

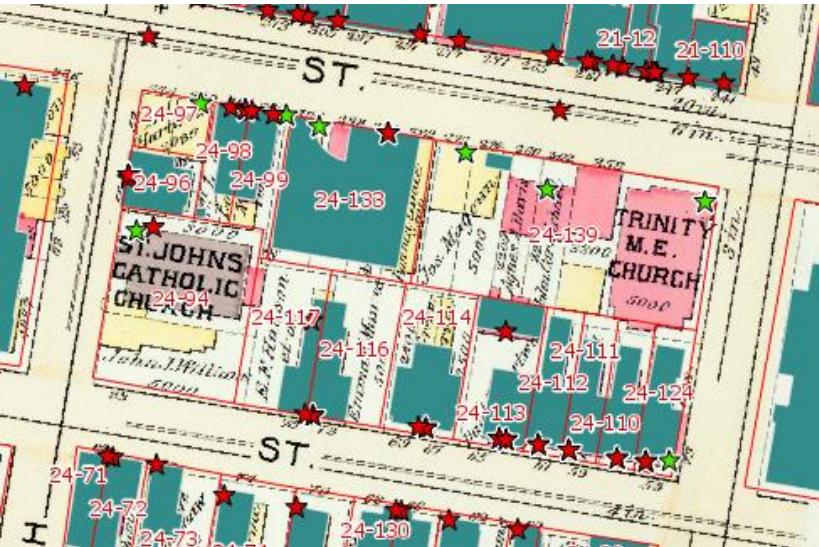
Result is a dictionary of all CHC Addresses with:

- Latitude
- Longitude,
- Map-Lot
- Building\_ID



- 256\_Cambridge\_St
- 262-266\_Cambridge\_St
- 276\_Cambridge\_St
- 292\_Cambridge\_St
- 304-306\_Cambridge\_St
- 308\_Cambridge\_St
- 314-316\_Cambridge\_St
- 320-325\_Cambridge\_St

	262-266_Cambridge_St_000-PE.jpg	JPG File	84.2 KB
	262-266_Cambridge_St_002p002-RF.jpg	JPG File	
	262-266_Cambridge_St_004p001-RN.jpg	JPG File	
	262-266_Cambridge_St_004p003-RN.jpg	JPG File	
	262-266_Cambridge_St_008-PL.jpg	JPG File	965 KB
	262-266_Cambridge_St_009p001-RN.jpg	JPG File	



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	42.36817
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
	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-delimited table formatted for the Omeka CSV import tool

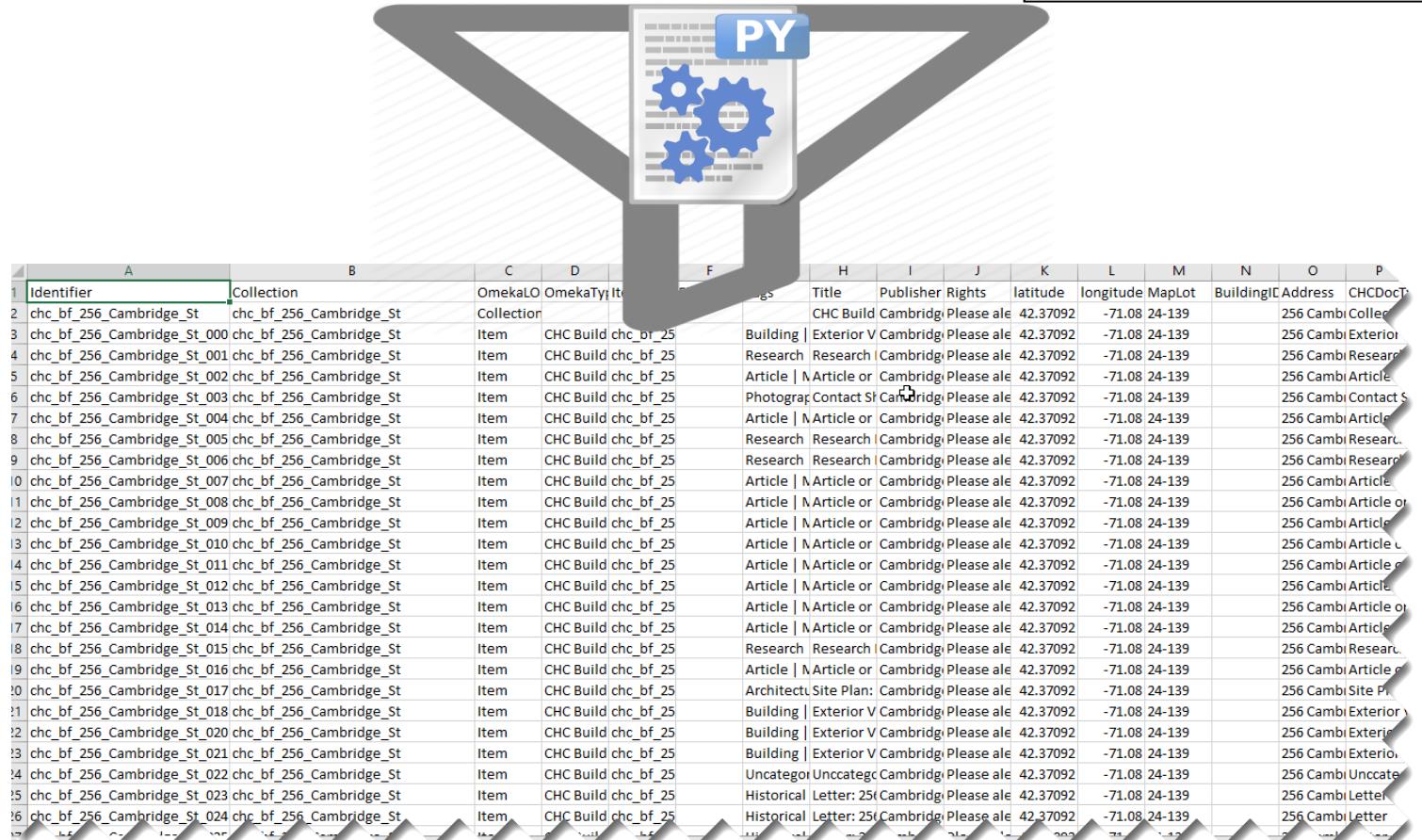
Future version could:

- 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	42.368117
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



# 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..

The screenshot shows the Omeka Admin interface for the 'CSV Import+' plugin. The URL in the browser is [40.71.40.46/omeka/admin/csv-import-plus/index/map-columns](http://40.71.40.46/omeka/admin/csv-import-plus/index/map-columns). The page title is 'Cambridge Historical Building Files Pilot'. The left sidebar menu includes 'Dashboard', 'Items', 'Collections', 'Item Types', 'Tags', 'Map', 'CSV Import+', 'Comments', and 'Exhibits'. The 'CSV Import+' item is highlighted. The main content area is titled 'CSV Import+'. It shows 'Step 2: Map columns to elements, tags, or files' for a CSV file named 'chc\_building\_files.csv'. A table maps CSV columns to Omeka elements:

Column header	Example from CSV File	Map To Element	Use HTML?	Special values	Extra data?
Identifier	chc_bf_256_Cambridge_St	Identifier	<input type="checkbox"/>	Identifier	<input type="checkbox"/>
Collection	chc_bf_256_Cambridge_St	Select Below	<input type="checkbox"/>	Collection (for item)	<input type="checkbox"/>
OmekaLOD	Collection	Select Below	<input type="checkbox"/>	Select Below	<input type="checkbox"/>
OmekaType		Select Below	<input type="checkbox"/>	Select Below	<input type="checkbox"/>
ItemID		Select Below	<input type="checkbox"/>	Select Below	<input type="checkbox"/>
File		Select Below	<input type="checkbox"/>	Files	<input type="checkbox"/>
Tags		Select Below	<input type="checkbox"/>	Tags	<input type="checkbox"/>
Title	CHC Building Files: 256 Cambridge St	Title	<input type="checkbox"/>	Select Below	<input type="checkbox"/>
Publisher	Cambridge Historical Commission	Publisher	<input type="checkbox"/>	Select Below	<input type="checkbox"/>
Rights	Please alert Cambridge Historical Commission nf	Rights	<input type="checkbox"/>	Select Below	<input type="checkbox"/>

# 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](http://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

The screenshot shows the Omeka.org homepage with a red header featuring the Omeka logo and navigation links for Showcase, Add-Ons, Forums, Documentation, Download, and Omeka S. Below the header are four green buttons: 'Serious Web Publishing', 'Cost-Effective Design', 'Flexible and Extensible', and 'Free and Open Source'. A large section titled 'Serious Web Publishing' describes creating complex narratives and sharing rich collections, with a 'Learn More' link. To the right is a 'Download Omeka' button for Linux, Apache, MySQL5, PHP5, and a 'System Requirements' link. Another button says 'No server? Try Omeka.net!'. The 'Omeka in Action' section includes a 'Tour Omeka' link and a screenshot of the Omeka interface showing item editing. The 'News' section features an article titled 'Omeka.net Refreshed and Renewed' dated 'August 28, 2017', which discusses the launch of a new look for Omeka.net and reduced prices for the Blue Plan.

# 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'"

Cambridge Historical Building Files Pilot

Search

Browse Items

Browse Collections

Map

Featured Item

No featured items are available.

Recently Added Items

Biographical Sketch: 73 Otis St

Site Plan: 73 Otis St

Featured Collection

No featured collections are available.

# Play with Public Omeka Site

Visit:

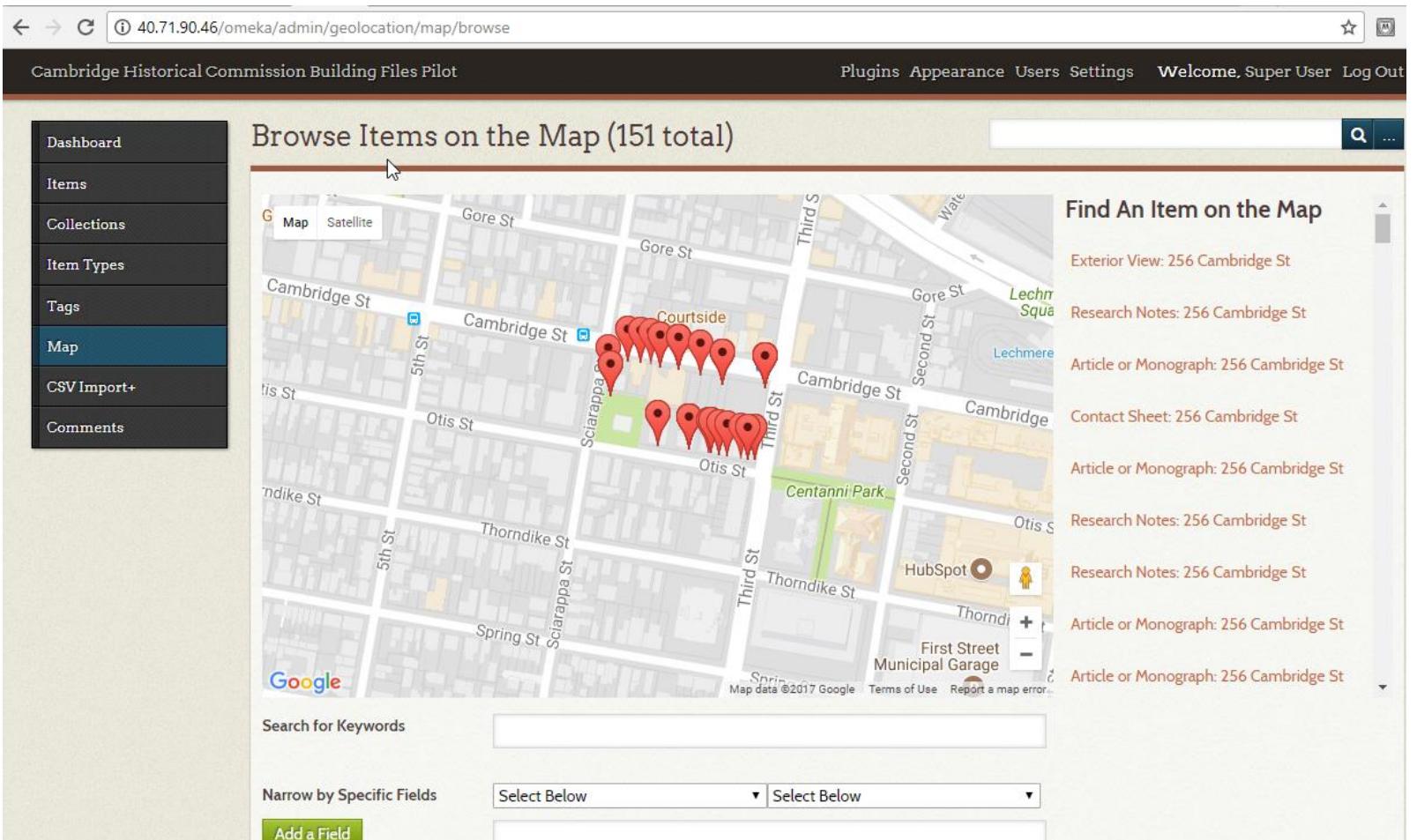
<http://40.71.90.46/omeka>

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.  
[www.pbcGIS.com](http://www.pbcGIS.com)



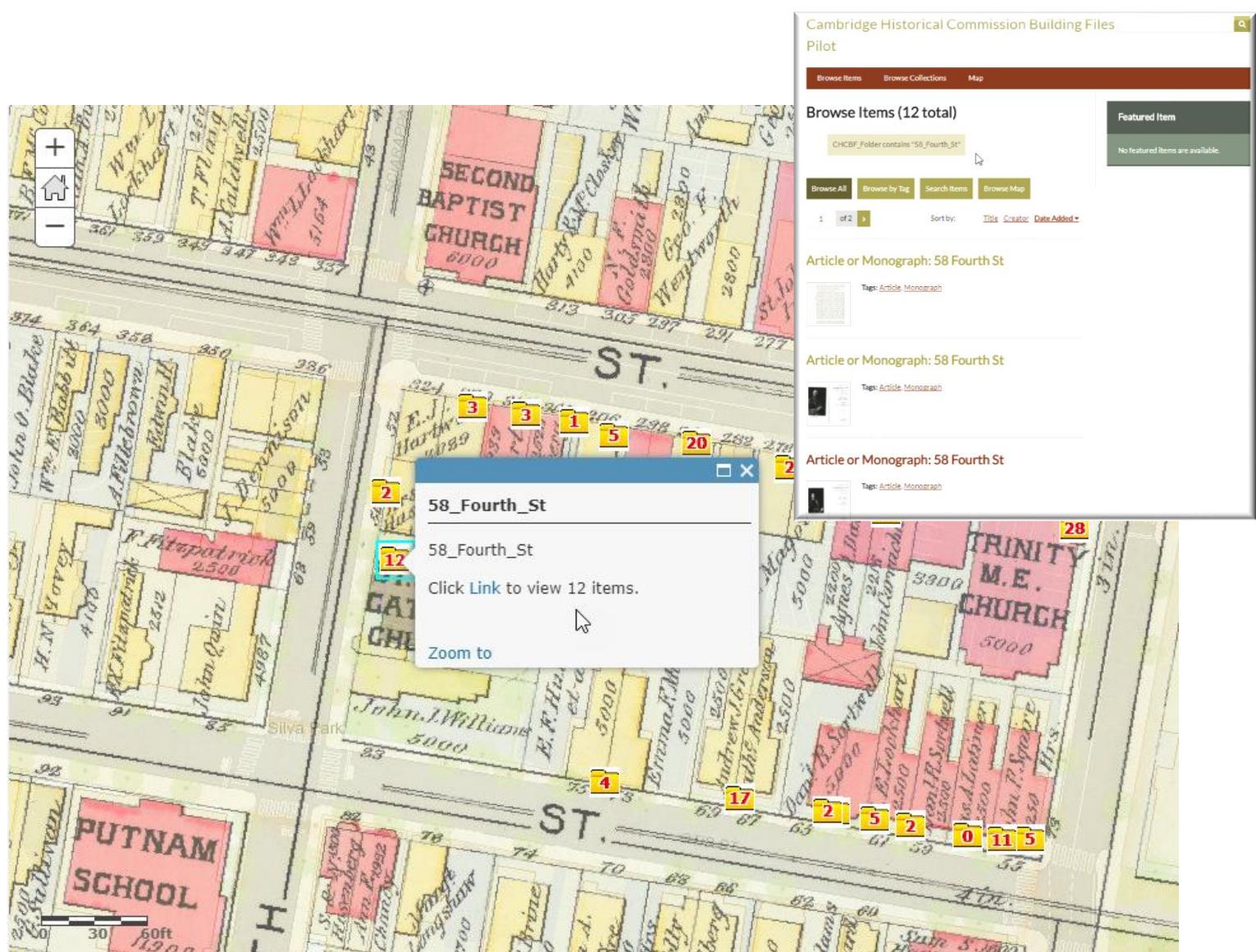
# Play with Pilot Web GIS Application

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 selected 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 Getty Vocabularies
- PDF Viewer / Export Plugin
- Neatline / Time Plugin
- CurateScape Plugin – make site like [ClevelandHistorical.org](http://ClevelandHistorical.org)

## Add and Manage Users

The screenshot shows the 'Appearance' section of the Omeka admin interface. A yellow box highlights the top navigation bar with links for Plugins, Appearance, Users, Setting, Welcome, Super User, and Log Out. Below this, the 'Themes' tab is selected. On the right, a preview of the 'Emiglio' theme is shown, featuring a dark header with red text and a light body with green accents. Below the preview, there's a 'Configure Theme' button and a note about adding themes from the Omeka Theme Directory or designing your own. Three other theme previews are shown at the bottom: 'Berlin' (blue header), 'Seasons' (light blue header), and 'Thanks, Roy' (white header).

Cambridge Historical Building Files Pilot

Plugins Appearance Users Setting

Welcome, Super User Log Out

Appearance

Themes Navigation Settings

Emiglio

Browse Items Browse Collections Browse Exhibits About Map

Aenean taciturna tibendum nulla sed consectetur. Nullam quis risus eget urna mollis ornare vel eu leo. Donec ullamcorper nulla non metus auctor fringilla. Fusce dapibus, tellus ac cursus commodo, tortor mauris condimentum nibh, ut fermentum massa justo sit amet. Vivamus sagittis lacus vel augue laoreet rutrum facilisis dolor auctor. Cras mattis consectetur purus sit amet fermentum.

Featured Item Rally to Restore Sanity and/or Fear

On October 30, 2010, John Stewart and Stephen Colbert had a rally on the National Mall. The event was a combination of Stewart's "Rally to Restore..."

Featured Collection Histories of the National Mall

Configure Theme

Add new themes by downloading them from the [Omeka Theme Directory](#), or design your own!

BERLIN

RECENTLY ADDED ITEMS George Washington to Eliza Kitchell of Washington

FEATURED ITEM Rally to Restore Sanity and/or Fear

Use this theme

Seasons

Featured Item

Featured Collection Histories of the National Mall

Featured Exhibit Rally to Restore Sanity and/or Fear

Use this theme

Thanks, Roy

George Washington to Eliza Kitchell of Washington

Featured Item Rally to Restore Sanity and/or Fear

Recently Added Items George Washington to Eliza Kitchell of Washington

Use this theme

Berlin

By Roy Rosenzweig Center for History and New Media

Seasons

By Roy Rosenzweig Center for History and New Media

Thanks, Roy

By Roy Rosenzweig Center for History and New Media

# 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](#)

The screenshot shows the Omeka Admin interface for the 'Cambridge Historical Building Files Pilot' site. The top navigation bar includes links for Plugins, Appearance, Users, Settings, Welcome, Super User, and Log Out. A search bar and a menu icon are also present.

The main content area displays item #520: "Biographical Sketch: 73 Otis St". On the left is a sidebar with links: Dashboard, Items (which is selected), Collections, Item Types, Tags, Map, CSV Import+, and Comments.

The central area contains Dublin Core metadata:

Title	Biographical Sketch: 73 Otis St
Subject	Otis
Publisher	Cambridge Historical Commission
Rights	Please alert Cambridge Historical Commission of any copyright questions or issues.
Identifier	chc_bf_73-75_Otis_St_004

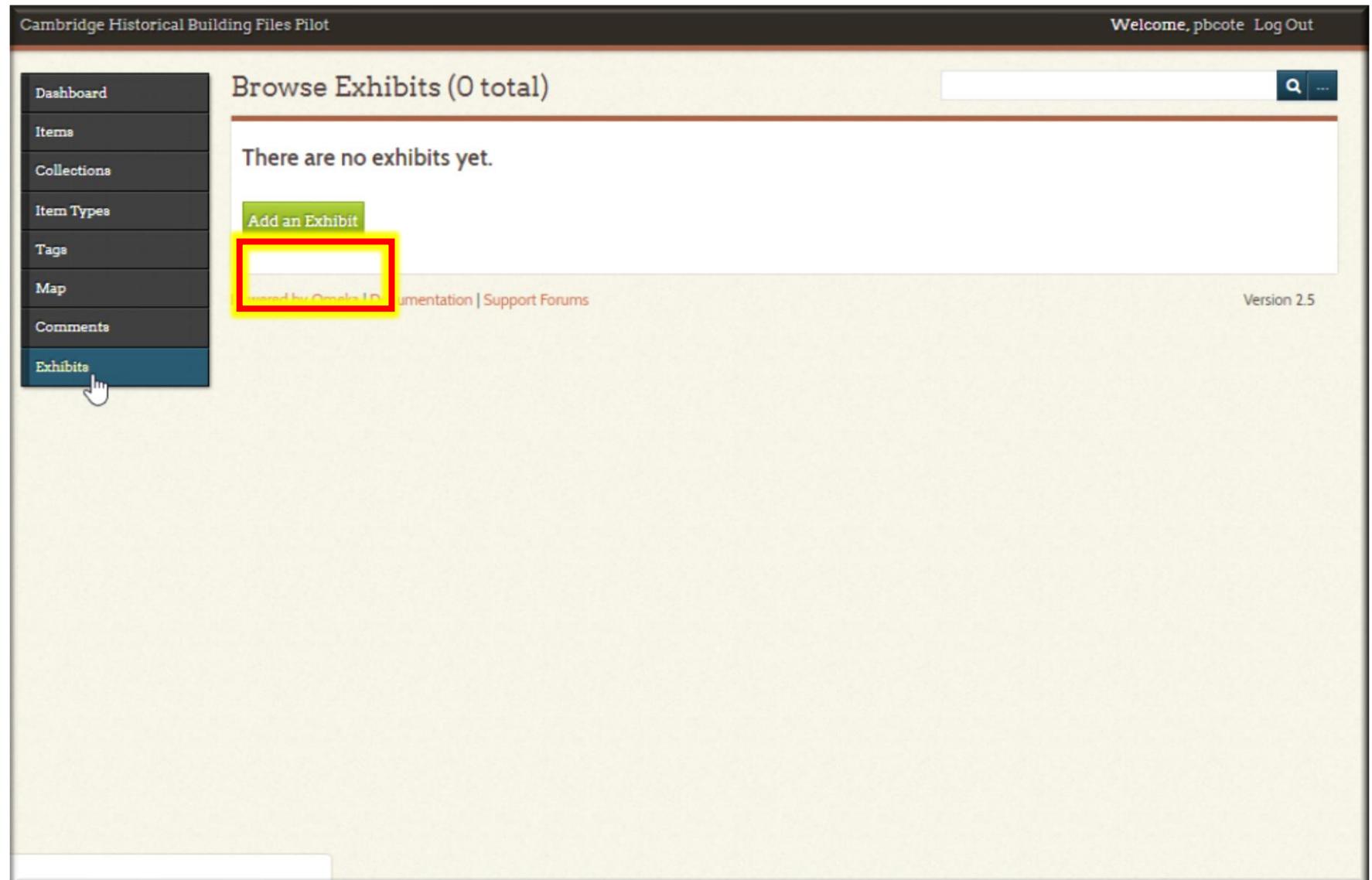
Below this is the 'CHC Building Files Item Type Metadata' section:

Address	73 Otis St
CHC Document Type	Uncategorized Item
Building ID	447-13
Map Lot	24-116
CHCBF_Folder	73-75_Otis_St

On the right side, there are several buttons: Prev Item, Edit (highlighted with a red box), View Public Page, and Delete. Below these are sections for Collection (CHC Building Files: 73-75 Otis St), Tags (Uncategorized), File Metadata (Uncategorized Item: 73 Otis St), and Output Formats (atom, dcmeta-xml, json, omeka-json).

# 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 Managment System

[Omeka Open Archive Initiative \(OAI\) Export Tool](#)

The screenshot shows the Omeka.org website interface. At the top is a red header bar with the Omeka logo and navigation links for Showcase, Add-Ons, Forums, Documentation, Download, and Omeka S. A search bar is also at the top. Below the header, the main content area has a yellow background. The title "Plugins/OaiPmhRepository 2.0" is displayed prominently. Below the title is a link to "Plugins". The main text describes the OaiPmhRepository plugin, mentioning its purpose of exposing metadata for Omeka items using the Open Archives Initiative Protocol for Metadata Harvesting. It also notes that this is the reverse of the functionality provided by the OaiPmhHarvester plugin. A section titled "Metadata Formats" follows, stating that the plugin ships with several default formats and can be altered or added. Three bullet points list the available formats: Dublin Core (oai\_dc), CDWA Lite (cdwalite), and MODS (mods). Each format has a corresponding bulleted list of details. To the right of the main content, there is a sidebar with sections for DOCUMENTATION (Home, Screencasts, Themes, Appearance, Plugins), PAGE VIEW (Page, Discussion, View source, History), and TOOLBOX (What links here, Related changes).

Plugins/OaiPmhRepository 2.0

< Plugins

The OaiPmhRepository plugin exposes metadata for Omeka items using the [Open Archives Initiative Protocol for Metadata Harvesting](#). This is the reverse of the functionality provided by the [OaiPmhHarvester plugin](#).

## Metadata Formats

The plugin ships with several default formats. Other plugins can alter these or add their own

- [Dublin Core](#) (oai\_dc)
  - This is required by the OAI-PMH specification for all repositories. Omeka metadata fields are mapped one-to-one with fields for this output format, and it is the preferred format to use with the plugin.
- [CDWA Lite](#) (cdwalite)
  - The mapping between Omeka's metadata and CDWA Lite metadata is more complicated, and certain fields may not be populated correctly. The chief advantage of using CDWA Lite output is that file URLs can be output in a controlled format, unlike Dublin Core. Harvesters may therefore be able to harvest or link to files in addition to metadata.
- [MODS](#) (mods)
  - This output crosswalks the Dublin Core metadata to MODS using the [mapping](#) recommended by the Library of Congress.

# Omeka Tour: Export Selected Items and Metadata

## HistoryPIN Upload Tool

The screenshot shows the HistoryPIN Bulk Uploader tool. At the top, there's a navigation bar with links for Home, Explore the Map, Meet our Members, Browse all Collections, News, Join, and Login. Below the navigation is the HistoryPIN logo. To the right of the logo are social media icons for Google+, Facebook, and Twitter. The main content area has a title 'Bulk Uploader for Chrome and Firefox'. Below the title is a bulleted list of questions: 'What is the Bulk Uploader?', 'How does it work?', 'What requirements are there?', and 'How to I get started?'. A large button labeled 'I'm ready to do a Bulk Upload' is prominently displayed. To the right of this section is a sidebar titled 'Downloadables' containing links for 'CSV template download', 'Instructions to complete the CSV', and 'Advice and tips'. At the bottom of the page, there's a section titled 'What is the Bulk Uploader?' with a detailed description and links for 'FAQs' and 'Street View overlaying'.

## Contribute Selected Items to Digital Commonwealth

The screenshot shows the Digital Commonwealth website. At the top, there's a navigation bar with links for Home, Search, Explore, For Libraries, For Educators, and About. The Digital Commonwealth logo is at the top left. To the right of the logo are fields for 'Email' and 'Password', a 'Remember me' checkbox, a 'Forgot password' link, and a 'Login' button. The main content area features a sidebar titled 'For Libraries' with sections for 'Membership', 'Mailing List Signup', 'Contribute Your Collections' (which is expanded to show 'Hosted Collections', 'Harvested Collections', and 'Metadata Requirements & Guidelines'), 'Events', 'Who We Are', and 'Contact Us'. The main content area has a heading 'Contribute Your Collections' with a sub-section titled 'Hosted Collections' which describes it as a solution for institutions with digitized materials. It also includes sections for 'Harvested Collections', 'Metadata Requirements & Guidelines', and a note about digitization services.

# 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>