

MESSAGE AGENCY

## LIBRARY SCIENCE

Integrating Drupal with DAMS

# A Quick Introduction



Who am I?

Aaron Bauman d.o user: aaronbauman

# Why Integrate?

- DAMS are usually powerful but ugly.
- Expose material in an interactive, curated user experience.
- You can appeal to a wider audience, by providing your content in multiple forms.
- Ease the content management burden when new material moves around on its own.

# What do I need?

- Someone who understands how to get stuff out of your DAMS.
- Someone who understands how to import things into Drupal.

Both of these are usually harder than they first appear.

# Case Studies

A black and white photograph of a modern library building with a grid of windows, situated in a courtyard with people and bicycles. The building is a large, rectangular structure with a prominent grid of windows. In the foreground, there is a paved courtyard with several bicycles parked in a row. People are walking around the courtyard. In the background, there are other buildings, including a large, ornate Gothic-style building with many windows. The sky is overcast.

Two examples:  
Beinecke Library at Yale  
Historical Society of Pennsylvania

# Beinecke Library

- Replace old static HTML site.
- Several DAMS that needed to be discoverable in one place.
- Provide rich media experience for great material.
- Ease the flow of new material from their DAMSs to the site.
- Keep the integration light weight.

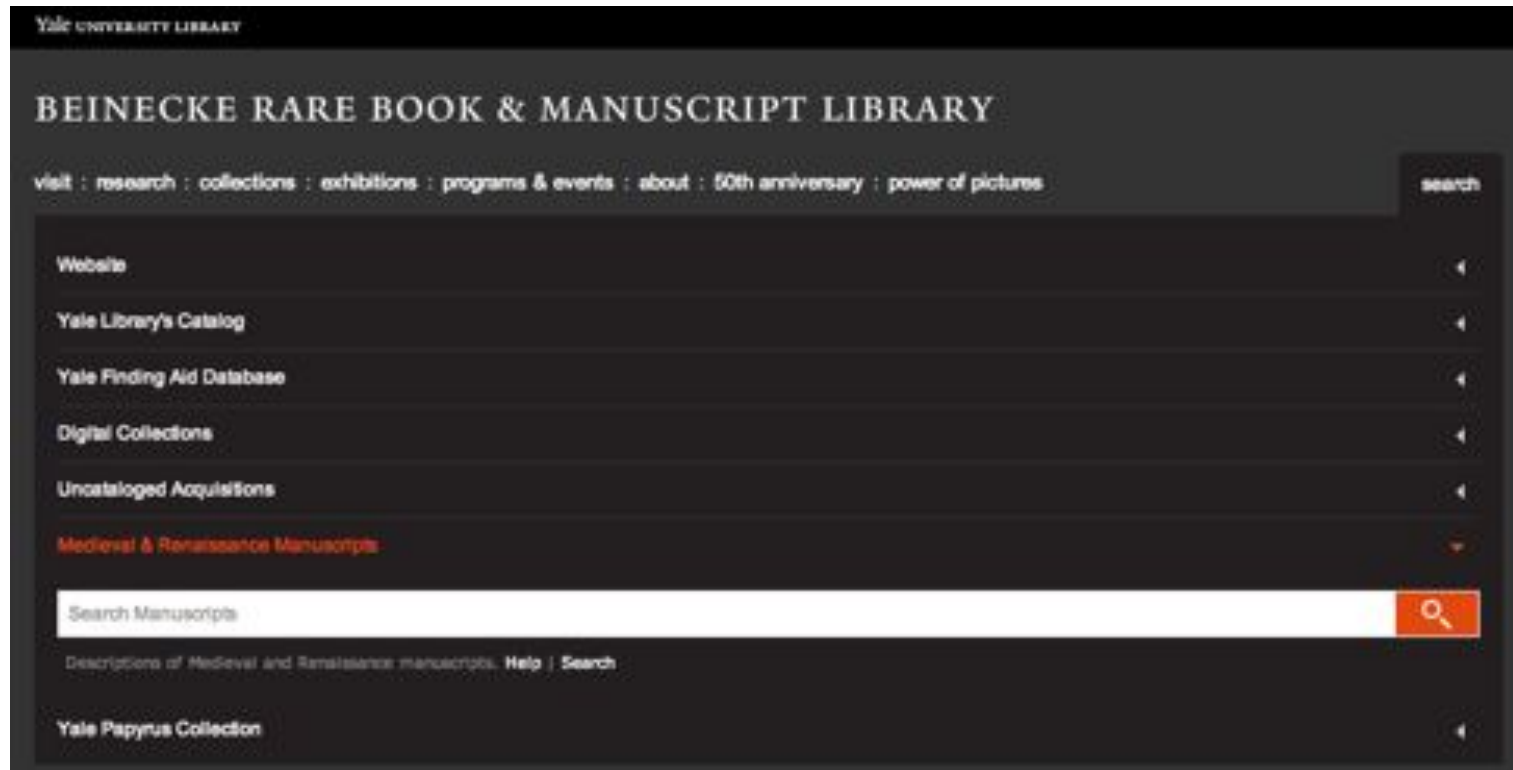


The Beinecke Rare Book & Manuscript Library is one of the world's largest libraries devoted entirely to rare books and manuscripts and is Yale's principal repository for literary archives, early manuscripts, and rare books. The Beinecke Library's robust collections are used to create new scholarship by researchers from around the world.

[Explore our Collection Highlights](#)

# Beinecke Library

## Level 1: Integrated Searches



The screenshot displays the Beinecke Rare Book & Manuscript Library website. At the top, it says "Yale UNIVERSITY LIBRARY" and "BEINECKE RARE BOOK & MANUSCRIPT LIBRARY". Below this is a navigation menu with links: "visit", "research", "collections", "exhibitions", "programs & events", "about", "50th anniversary", and "power of pictures". A "search" button is located in the top right corner. The main content area features a list of links: "WebSite", "Yale Library's Catalog", "Yale Finding Aid Database", "Digital Collections", "Uncataloged Acquisitions", and "Medieval & Renaissance Manuscripts". The "Medieval & Renaissance Manuscripts" link is highlighted in red. Below this list is a search bar with the text "Search Manuscripts" and a magnifying glass icon. At the bottom of the search bar, it says "Descriptions of Medieval and Renaissance manuscripts. Help | Search". The "Yale Papyrus Collection" link is visible at the bottom of the page.

Yale UNIVERSITY LIBRARY

BEINECKE RARE BOOK & MANUSCRIPT LIBRARY

visit : research : collections : exhibitions : programs & events : about : 50th anniversary : power of pictures search

WebSite

Yale Library's Catalog

Yale Finding Aid Database

Digital Collections

Uncataloged Acquisitions

Medieval & Renaissance Manuscripts

Search Manuscripts

Descriptions of Medieval and Renaissance manuscripts. Help | Search

Yale Papyrus Collection

# Beinecke Library

## Level 2: XML and RSS import

Yale UNIVERSITY LIBRARY

### BEINECKE RARE BOOK & MANUSCRIPT LIBRARY

[visit](#) : [research](#) : [collections](#) : [exhibitions](#) : [programs & events](#) : [about](#) : [50th anniversary](#) : [power of pictures](#)

search ▾

## Import

Import	Description
<a href="#">New Acquisitions</a>	Imports the list of new acquisitions in the system.
<a href="#">New Finding Aids</a>	Import new finding aids.



# Beinecke Library

## Level 3: Relate content and export

Home » Administration » Structure » Content types » Exhibition

### Exhibition

LABEL	MACHINE NAME
↕ Title	title
↕ Subtitle	field_subtitle
↕ Signature Image	field_signature_image
↕ Exhibition Type	field_exhibition_type
↕ Overview	body
↕ Date	field_date
↕ Location	field_location
↕ Sections	field_sections
↕ Landing Page Box	field_landing_page_box
↕ Extra Exhibition Links	field_extra_exhibit_links
↕ Library ID	field_library_id
↕ IBI with caption	ibic

Home » Administration » Structure

### Services

Services are collections of methods available to remote applications. They are defined in modules, and information.

All enabled services and methods are shown. Click on any method to view information or test.

[+ Add](#) [+ Import](#)

Storage

- All -

Enabled

- All -

Search

Sort by

Enabled, name

Order

Up

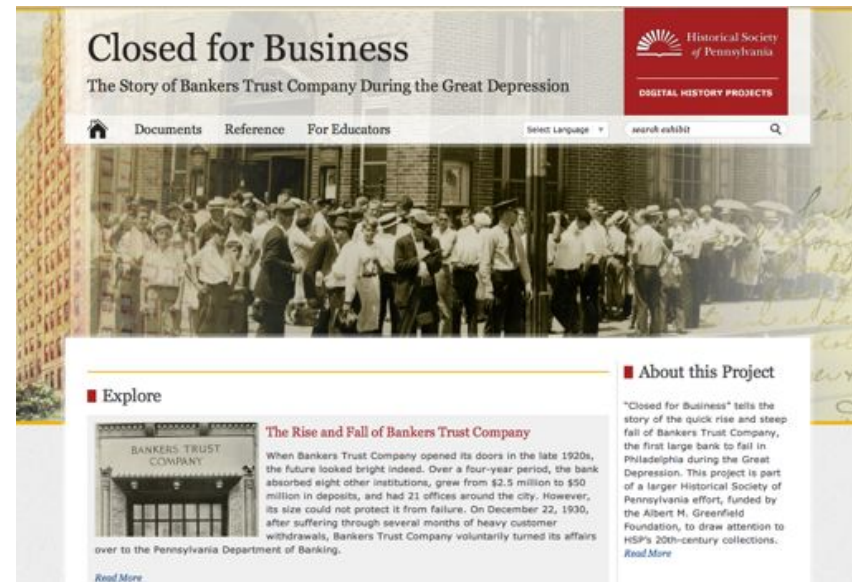
Apply

Reset

NAME	STORAGE
id_search	Normal

# Historical Society of Pennsylvania

- Digital History Exhibits Project
- Client need: Branded digital exhibit template
- Drupal content generated by importing a TEI-encoded XML document
- Pulls digital assets from DAMS via API



# Historical Society of Pennsylvania

## Exhibit Structure

### → Objectives

- ◆ Drupal admins can create new exhibits as needed
- ◆ Curated narrative to expose depth of content
- ◆ Relate content to primary asset (documents)

### → Application

- ◆ Node type drives context via Reference fields
- ◆ Feeds, XML Parser
- ◆ Customizations to accommodate TEI, field collections, other exigencies

# Historical Society of Pennsylvania

Home + Add content

## Create Exhibit

Exhibit Title \*

Subtitle

Banner image

Add

• INTRODUCTION

Header

Introduction Text

Text format TECH ⌵ [More information about text formats](#)

• EXHIBIT CONTENT

[Show row weights](#)

SECTIONS:

+

+ Create Exhibit Section 🔍 Search

Add another item

# Historical Society of Pennsylvania

Closed for Business | exhibits.hsp.org

digitalhistory.hsp.org/closed-business

## Closed for Business

The Story of Bankers Trust Company During the Great Depression

Home Documents Reference For Educators Select Language



■ Explore

The Rise and Fall of Bankers Trust Company

The image shows a screenshot of a digital exhibit page. At the top, the browser address bar shows 'digitalhistory.hsp.org/closed-business'. The main heading is 'Closed for Business' in a large, bold, serif font, with the subtitle 'The Story of Bankers Trust Company During the Great Depression' below it. A navigation bar contains a home icon, 'Documents', 'Reference', 'For Educators', and a 'Select Language' link. Below the navigation is a large black and white photograph of a dense crowd of people, many wearing hats, gathered outside a building. At the bottom, there is an 'Explore' section with a small thumbnail image and the text 'The Rise and Fall of Bankers Trust Company'. Three red arrows point from the left side of the image towards the main heading and the photograph.

# Historical Society of Pennsylvania

Preserving American Freedom | exhibits.hsp.org

digitalhistory.hsp.org/preserving-american-freedom

## Preserving American Freedom

The Evolution of American Liberties in Fifty Documents

Home Documents Reference Timeline For Educators Select Language

IN CONGRESS, JULY 4, 1776  
A DECLARATION  
BY THE REPRESENTATIVES OF THE  
UNITED STATES OF AMERICA  
IN GENERAL CONGRESS ASSEMBLED

Introduction 1 of 8

### The Contested History of American Freedom

A thematic essay by Pulitzer prize-winning historian Eric

# Historical Society of Pennsylvania

## TEI Parser

### → Objectives

- ◆ Semantic encoding of historical documents
- ◆ Consolidation and unification: D.R.Y.

### → Application

- ◆ Import TEI (xml) file to create/update Drupal nodes
  - Feeds
  - XPath Parser
  - hooks
  - patches
- ◆ **Warning:** feeds and field collections don't get along

# Historical Society of Pennsylvania

## Notable Hooks and patches

- `views_query_alter`
- `block_view_alter`
- `menu_block_tree_alter`
- `search_api_page_search_form_alter`
- `search_api_query_alter`
- `menu_alter`
- `form_feeds_import_form_alter`
- `feeds_after_parse`
- `feeds_presave`

- Patches created by Message Agency are always contributed back to drupal.org
- `1424086--ignore_destination_GET_parameter-1.patch`
- `2006870-1--negated_fulltext_filters.patch`
- `date-date_plain_format_settings-1806856-2.patch`
- `date-plain_text_formatter_respect_settings-1763788-3.patch`
- `eck-admin_paths-1836892-1.patch`
- `entityreference-feeds_integration-1616680-66.patch`
- `feeds-DateTime_support_for_php_5.3-1258582.patch`
- `field_collection-feeds-1063434-93.patch`
- `field_collection_feeds_multiple.patch`
- `hide-empty-field-collections-1276258-33_0.patch`
- `leaflet-markercluterer-outsideInteraction-100644.diff`
- `relation-7.x-1.x-dev--feeds-integration-1345958.patch`
- `search_api-1694832-7--field_changes.patch`
- `search_api-add_templates-1313676-26.patch`
- `search_api_db-prevent_infinite_recursion-1917592_0.patch`
- `search_api_facetapi_0.patch`
- `timeline-date-field.patch`



# Historical Society of Pennsylvania

## Closed for Business

The Story of Bankers Trust Company During the Great Depression



[Documents](#)

[Reference](#)

[For Educators](#)

Select Language ▾



### Import

### Description

[Single Document](#)

ONE-FILE-AT-A-TIME Single Document importer

[Multi-Document Corpus](#)

MULTIPLE file upload for documents containing multiple transcripts

[Genres](#)

[Organizations](#)

[People](#)

[Topics](#)

# Historical Society of Pennsylvania

Import

Delete items

Log

Unlock

Status

Last import: 3 weeks ago,  
357 imported items total.

Import

**Related Exhibit** \*

- Select -

File

8748-2000B-103f4.xml

22.25 KB

application/xml

Choose file No file chosen

Select a different file from your local system.

Import

# Historical Society of Pennsylvania

Show row weights

## DOCUMENT DETAILS:

### Source Image \*

80:5264-1958, b14092, 1.tif

### Document Title \*

Telegram from Irland Beckman to Albert Greenfield

### DATE (SEARCH)

Show End Date

1900/12/22

Format: 2013/10/13

### User-Facing Date

December 22, 1900

Show row weights

## AUTHOR(S): \*



Irland McKnight Beckman (4027)



Add another item

Show row weights

## RECIPIENT(S):



Albert Monroe Greenfield (4253)



Add another item

# Historical Society of Pennsylvania

## XSL Filter

### → Objectives

- ◆ Non-destructive import of TEI markup
- ◆ Provide source XML documents for download

### → Application

- ◆ TEI CHI input filter provides starting point
- ◆ Use XSL to transform TEI markup into HTML

# Historical Society of Pennsylvania

## XSL Filter

### → Key features

- ◆ Tailored XSLT and XSL stylesheet
- ◆ “Fuzzy dates” for display; real dates for search
- ◆ XML id -> Feeds GUID -> Node id translation
- ◆ Modals!

# Historical Society of Pennsylvania

## Transcript

```
<p>PATRONS ARE REQUESTED TO FAVOR THE COMPANY BY CRITICISM AND SUGGESTION CONCERNING ITS  
SERVICE</p>  
<p>12015</p>  
<p>CLASS OF SERVICE</p>  
<p>This is a full-rate Telegram or Cablegram unless its deferred character is indicated by a  
suitable sign above or preceding the address.</p>  
<p><orgName ref="Greenfield_project_organizations.xml#Western_Union">WESTERN  
UNION</orgName></p>  
<p><persName ref="Greenfield_project_people.xml#Carlton_Newcomb">NEWCOMB CARLTON</persName>
```



Text format

TEIHL



[More information about text formats](#)

## Source Information

```
<sourceDesc>  
<msDesc>  
<msIdentifier>  
<repository>The Historical Society of Pennsylvania, Philadelphia, PA</repository>  
<collection>Albert M. Greenfield papers (Collection 1958)</collection>
```

Text format

TEIHL



[More information about text formats](#)

## Source Page

## Source ID

Remove

# Historical Society of Pennsylvania

Show row weights

## DOCUMENT DETAILS:

### Source Image \*

80:5264-1958, b14092, 1.tif

### Document Title \*

Telegram from Irland Beckman to Albert Greenfield

### DATE (SEARCH)

Show End Date

1900/12/22

Format: 2013/10/11



### User-Facing Date

December 22, 1900



Show row weights

## AUTHOR(S): \*



Irland McKnight Beckman (4027)



Add another item

Show row weights

## RECIPIENT(S):



Albert Monroe Greenfield (4253)



Add another item

# Historical Society of Pennsylvania

## Document Viewer

### → Objectives

- ◆ Expose DAMS assets to Drupal, via TEI
- ◆ Page and transcript side-by-side
- ◆ Access to high resolution image for pan and zoom

### → Application

- ◆ Collective Access API
- ◆ Asynchronous image caching
- ◆ New York Times Viewer (OSS, via github)



# Historical Society of Pennsylvania

Show row weights

## DOCUMENT DETAILS:

### Source Image \*

80:5284-1888, D14092, 1 MF

### Document Title \*

Telegram from Irland Beckman to Albert Greenfield

### DATE (SEARCH)

Show End Date

1900/12/22

Format: 2013/10/11

### User-Facing Date

December 22, 1900

Show row weights

## AUTHOR(S): \*



Irland McKnight Beckman (4027)



Add another item

Show row weights

## RECIPIENT(S):



Albert Monroe Greenfield (4253)




Add another item

# Historical Society of Pennsylvania

## Telegram from Irland Beckman to Albert Greenfield, December 22, 1930

Telegram from Irland Beckman to Albert Greenfield 1 of 1

Author(s): Irland McKnight Beckman | Recipient(s): Albert Monroe Greenfield |  
December 22, 1930



PATRONS ARE REQUESTED TO FAVOR THE COMPANY BY CRITICISM AND SUGGESTION CONCERNING ITS SERVICE

12015

CLASS OF SERVICE

This is a full-rate Telegram or Cablegram unless its deferred character is indicated by a suitable sign above or preceding the address.

WESTERN UNION

NEWCOMB CARLTON, PRESIDENT


J. C. WILLEVER, FIRST VICE-PRESIDENT

Transcript

Source Image

Entity Reference

# Historical Society of Pennsylvania



**WESTERN UNION**

11 D&A 55 19 3 EXTRA RUSH A  
PHILADELPHIA PENN DEC 22 400A  
1930 ALBERT M GREENFIELD  
DELIVER IMMEDIATELY GET  
RECEIPT 310 WEST JOHNSON ST  
PHILA

SPECIAL MEETING OF BOARD THIS  
MORNING DECEMBER 22ND AT  
EIGHT OCLOCK YOUR PRESENCE  
IMPERATIVE

IRLAND M BECKMAN  
SECRETARY BOARD 420 A  
[shorthand text]

**Note 1**  
This meeting no dou  
that Bankers Trust C  
for business that day  
turned over its affair  
of Banking.

**Modal!**

**Source Information:**  
**Title**  
Telegram from Irland Beckman to Albert Greenfield

# Historical Society of Pennsylvania

Telegram from Irland Beckman to Albert Greenfield

Author(s): Irland Beckman

Recipient(s): Albert M. Greenfield

December 22, 1930

**Modal!**

Albert Monroe Greenfield



Albert M. Greenfield was a prominent Philadelphia businessman involved in real estate, banking, transportation, and retail, as well as philanthropy, etc. etc. etc. and etc. etc. etc.



WESTERN UNION GIFT ORDERS  
ARE APPROPRIATE GIFTS FOR ALL  
OCCASIONS.

# Historical Society of Pennsylvania

Telegram from Irland Beckman to Albert Greenfield, December 22, 1930 | exhibits.hsp.org

digitalhistory.hsp.org /node/7279

DOCUMENT NOTES

PHILADELPHIA PENN DEC 22 400A 1930 ALBERT M GREENFIELD DELIVER IMMEDIATELY GET RECEIPT 350 WEST

THIS MORNING DECEMBER 22ND AT EIGHT OCLOCK YOUR PRESENCE IMPERATIVE

[shorthand text]

RS ARE APPROPRIATE GIFTS FOR ALL OCCASIONS.

WE REQUESTED TO FAVOR THE COMPANY BY CRITICISM AND SUGGESTION CONCERNING ITS SERVICE

# WESTERN UNION

SIGNS
DL = Day Letter
NM = Night Message
NL = Night Letter
LCO = Deferred Call
NLT = Cable Night Letter
WLT = Week-End Letter

NEWCOMB CARLTON, PRESIDENT A. C. WILLEYER, FIRST VICE-PRESIDENT

be date time on full-rate telegrams and day letters, and the time of receipt at destination as shown on all messages, is STANDARD TIME

ern Union Building, 230 So. 11th St., Philadelphia, Pa.

ARE APPROPRIATE GIFTS FOR ALL

Zoom

Page 1 of 1

- Telegram from Irland Beckman to Albert Greenfield
- Transcript

Modal!

# Historical Society of Pennsylvania

## Event Timeline and Map

### → Objectives

- ◆ Personal narratives of the Underground Railroad
- ◆ Extract normalized data from historic documents
- ◆ Expose stories via spatial relationships
- ◆ Navigate spatiotemporal data

### → Application

- ◆ Geocode location data
- ◆ Views-driven, contextual map + timeline
- ◆ Time Mapper: <http://timemapper.okfnlabs.org>

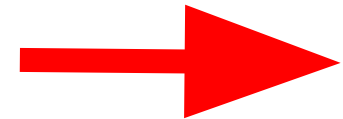
# Historical Society of Pennsylvania





# Historical Society of Pennsylvania

The screenshot displays the 'Underground Railroad' website interface. At the top, there is a navigation menu with 'Documents', 'Reference', 'Places', 'Relationships', and 'Educator Tools'. Below the menu is a header image of a group of people. The main content area is titled 'Locations on the Underground Railroad' and includes a map of the United States with several location markers. A red box highlights the map area with the word 'Map' in large red text. Below the map, there is a section titled 'Canada' with a timeline and a red box highlighting it with the word 'Timeline' in large red text. The website footer includes the 'Historical Society of Pennsylvania' logo and an 'E-NEWS' button.





# Historical Society of Pennsylvania

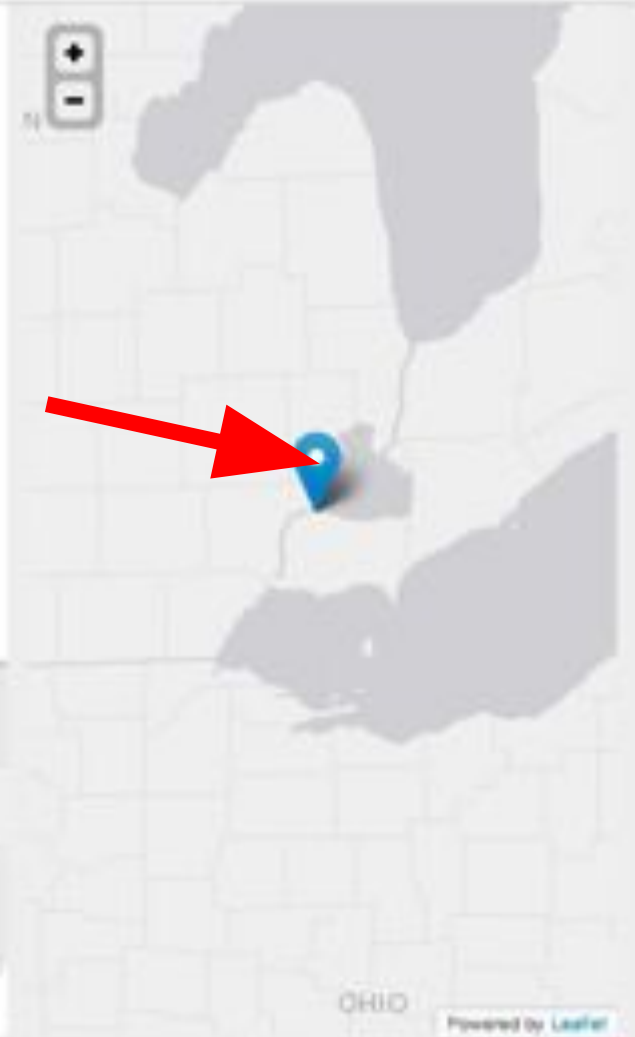
## Events at this location

circa late 1855 - early 1856

### Canada

The Vigilance Committee in Philadelphia intends to send Harriet Shephard's group to Canada, but divides and disguises the group for safety.

Related: [Anna Maria Shephard](#), [Edwin Shephard](#), [Eliza Jane Shephard](#), [Harriet Shephard](#), [John Henry Shephard](#), [Mary Ann Shephard](#), [Unnamed Fugitive Man 1](#), [Unnamed Fugitive Woman 2](#), [Hiram Wilson](#)



# Historical Society of Pennsylvania

## Relationship Network

### → Objectives

- ◆ Surface relationships between all individuals
- ◆ Discover buried stories within network
- ◆ Visualize interactions between key players

### → Application

- ◆ D3.js Force Directed network graph <http://d3js.org>
- ◆ Again, views-driven, contextual graph
- ◆ MIT Media Lab “Immersion” project: <https://immersion.media.mit.edu/demo>

# Historical Society of Pennsylvania



# Historical Society of Pennsylvania



# Historical Society of Pennsylvania



# Historical Society of Pennsylvania

## Lessons Learned

- The data will not compile itself
- Planning is important
- Flexibility is more important
- The data may not actually tell a good story
- Limits of drupal contrib
- Decide on a vendor, not a pricetag

# Historical Society of Pennsylvania

## The data will not compile itself

- ◆ Massive volume of man hours parsing historic documents, organizing and normalizing.
- ◆ Underestimating or underfunding this effort will impact a project timeline.
- ◆ Once the data is extracted, make sure you have a good place to put it.

# Historical Society of Pennsylvania

## Planning is important

- ◆ What data is worth collecting?
- ◆ How granular can the data be?
- ◆ What DAMS suits our needs?
- ◆ What goals should the application fulfill?
- ◆ What should be prioritized, and what can slip?
- ◆ What is the Minimum Viable Product?
  
- ◆ Bring vendors, consultants, collaborators in early.
- ◆ But not too early.



# Historical Society of Pennsylvania

## Flexibility is more important

- ◆ Planning the data schema is necessary to integrate with 3rd party systems, but some degree of schema shift should be expected.
- ◆ Plan an extensible, pluggable system.
- ◆ The lowest bidder is never planning for the next project.
- ◆ Don't hack core.
- ◆ Don't hack contrib.
- ◆ Do not reinvent the wheel.

# Historical Society of Pennsylvania

## Data may not actually tell a good story

- ◆ You may be expecting too much from your data.
- ◆ Data curation can be just as effort intensive as data collection.
- ◆ Data by itself is not necessarily interesting.
- ◆ Think about your target audiences and user stories.
- ◆ Design and presentation are just as important as content in order to reach your constituents.

# Historical Society of Pennsylvania

## Limits of Drupal contrib

- ◆ Please refer to slide titled “Notable Hooks and patches”
- ◆ Just because there’s a module called “Disqus” doesn’t mean it does what you think it does.
- ◆ Just because you know how two modules work doesn’t mean you know how they work together.
- ◆ Conflicting pressures: “Right Way” vs. “Fast Way”

# Historical Society of Pennsylvania

## Decide on a vendor, not a pricetag

- ◆ Estimation and RFPs are guess work.
- ◆ Specify budget or features. Not both.
- ◆ Mantra: good, fast, cheap. Choose 2.
- ◆ Having a good relationship with a talented vendor will save you time, money, and headache. Full stop.

# Questions?



# Resources

- D3.js  
<http://d3js.org>
- Immersion  
<http://immersion.media.mit.edu>
- TimeMapper  
<http://timemapper.okfnlabs.org>
- NYT Viewer  
<https://github.com/documentcloud/document-viewer>
- Leaflet  
<http://leafletjs.com>
- TEI  
<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>
- TEICHI  
<http://www.teichi.org/>
- Collective Access  
<http://www.collectiveaccess.org>

# MESSAGE AGENCY

Aaron Bauman

[aaron@messageagency.com](mailto:aaron@messageagency.com)

<http://drupal.org/user/384578>

Salesforce Suite Project Owner & Maintainer

<http://messageagency.com>

[info@messageagency.com](mailto:info@messageagency.com)

(215) 546-6496