



Editing Women's Papers for Publication on the Web

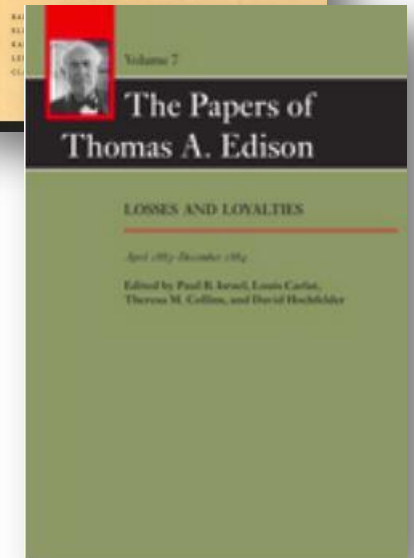
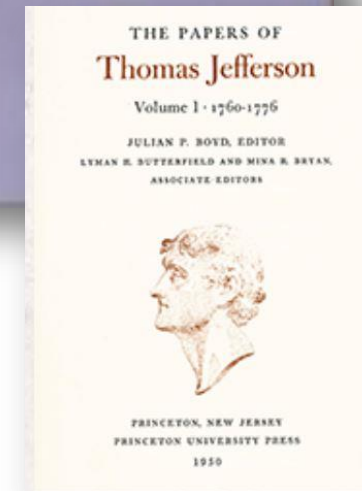
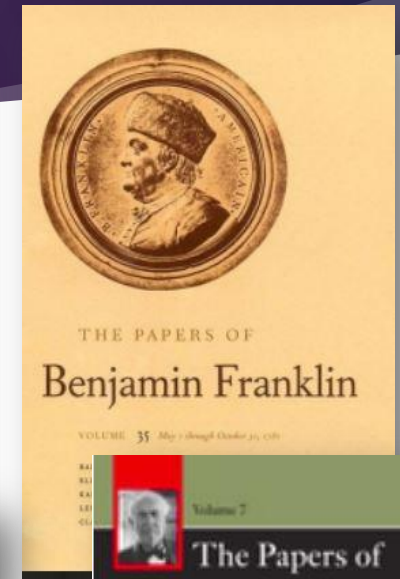
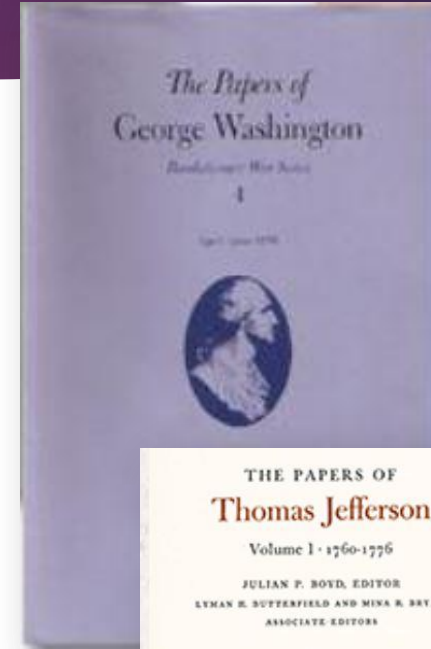
CATHY MORAN HAJO

RAMAPO COLLEGE OF NEW JERSEY

Documentary, Scholarly, or Historical, Editing

Small community of historians who work to publish primary source materials.

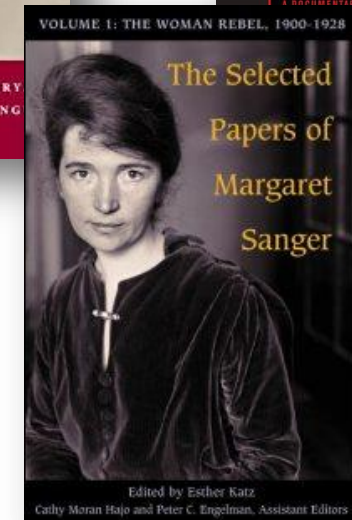
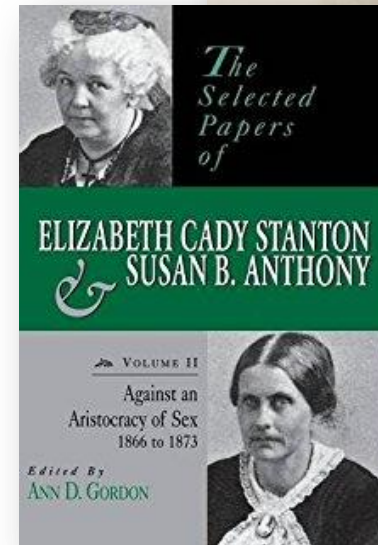
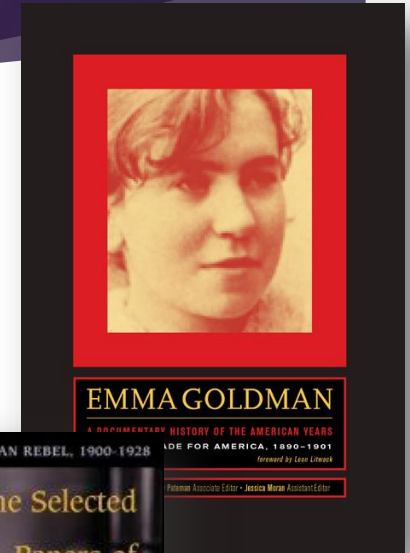
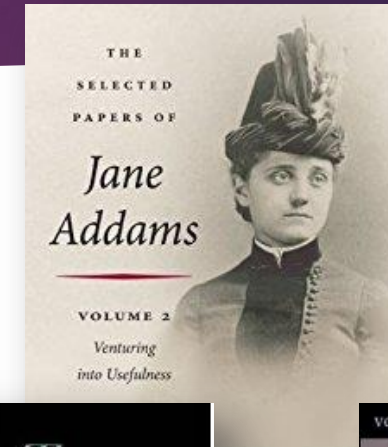
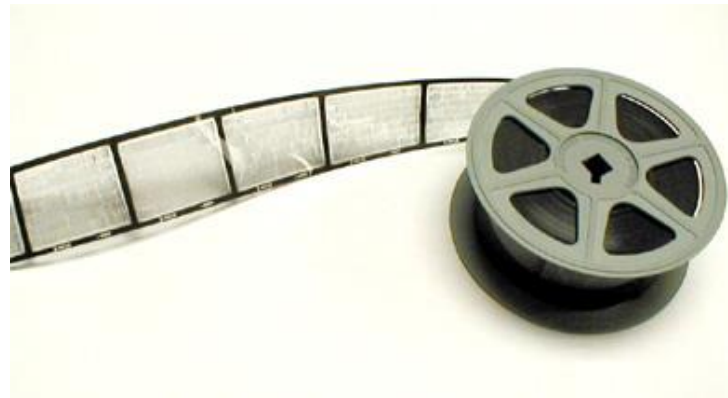
- Earliest editions of Founding Fathers begun in 1950s with slow expansion to include broader figures.
- Most published multi-volume comprehensive editions of “great white men.”
- Include transcriptions of primary sources with annotation that helps contextualize the documents.
- Teams work for years to produce these projects whose content should last for generations.



Women's Projects

In the 1970s and 1980s women's projects began

- ▶ Limited number of projects began.
- ▶ No comprehensive books; instead the “microfilm ghetto” and short, highly selected book editions.



Digitizing Editions

- ▶ Early efforts focused on digitizing the already published Founding Era volumes.
- ▶ Published by the University of Virginia's electronic imprint, Rotunda as *The American Founding Era*.
- ▶ As new volumes were published, they are added to the digital edition.
- ▶ Free limited version, *Founder's Online*, created later.
- ▶ Not the best model going forward.



Two Editions of Women's Papers



▶ **Margaret Sanger Papers**

- ▶ Founded in 1987
- ▶ 245 reel microfilm of 95,000 documents
- ▶ 4 volume book edition of about 1,000 documents.
- ▶ Years covered 1900-1966

▶ **Jane Addams Papers**

- ▶ Founded in 1975
- ▶ 82-reel microfilm
- ▶ 6 volume book edition planned, 2 published
- ▶ Years covered 1860-1935



Challenges

▶ Copyright

- ▶ Public domain
- ▶ Fair Use

▶ Scale

- ▶ Number of documents
- ▶ Whether you will include all
- ▶ Variety of materials.

▶ Funding

- ▶ Institution
- ▶ Federal Grants
- ▶ Private foundations and donations

▶ Technology

- ▶ Options
- ▶ Institutional Support
- ▶ Open source

Margaret Sanger Papers

- ▶ Began in 2003.
- ▶ Includes about 1,200 documents in transcribed form.
- ▶ Each transcription is encoded using XML to identify parts of the document, regularize them, and then search for them.
- ▶ Currently quite unattractive!

THE *Margaret Sanger* PAPERS PROJECT
PUBLISHED WRITINGS AND SPEECHES, 1911-1959



XML - Extensible Markup Language

The Good

- Extremely flexible and portable
- Extremely powerful
 - Can encode multiple versions of a text
 - Can be reused easily
 - Allows very detailed searches

The Bad

- Sharp learning curve.
- Difficult to work with when editing and proofreading
- Difficult to publish without programming skills.

And it is ugly.

```
<docBody>
  <head>Reception for opening of Clinic</head>
  <byline>M.S. remarks</byline>
  <p><person reg="Kato, Shidzue Ishimoto">Baroness Ishimoto</person>, and friends:</p>
  <p>I am very happy indeed to be in <place>Japan</place> again and to meet so many old
    friends. Some of them have been to <place reg="United States">America</place> and have
    visited us at our large <org reg="Birth Control Clinical Research Bureau">Demonstration
      Clinic</org> in <place reg="New York, NY">New York</place>, and other I met in 1922
    in Japan on my first visit. <del>So I feel</del> You all never seem to be any older.</p>
  <p>I, however, seem to be unfortunate in my visits to Japan. The first time I came with
    enthusiasm. I felt I had a message for the mothers of your country. I was not so
    hospitably received by your government. Nevertheless, the people, the mothers and the
    plain people, welcome me and gave me inspiration to carry this message on not only in
    Japan but throughout the Orient.</p>
  <p>On my second visit Japan was under martial law, and no meetings were allowed. This time
    on my way to <place>China</place> and <place reg="Manila, Philippines">Manila</place> I
    fell and broke my arm, and in view of the difficulty between China and Japan, I have
    decided to postpone my visit to China to another year.</p>
  <p>Every country has its pioneers. <place>England</place> had its <person
    reg="Wollstonecraft, Mary">Mary Wollencraft</person>, <place>France</place> its
    <person reg="Curie, Marie">Madame Curie</person>, America its <person
    reg="Anthony, Susan B.">Susan B. Anthony</person>. Many others, <person
    reg="Schreiner, Oliver">Oliver Shriner</person>, <person reg="Key, Ellen">Ellen
    Key</person> of the <place reg="Scandinavia">Scandinavian countries</place>; every
    country has men who pioneer in their line and women in others, and Japan should be
    proud, and I believe her children's children will pay homage to one of the greatest
    women today, to Baroness Ishimoto, for Baroness Ishimoto is not only a pioneer in Japan,
    but also a <person reg="Polo, Marco">Marco Polo</person> of Japan, because thru her
```


XML and Editing

- ▶ Looks like HTML, but it is doing something different.
- ▶ XML gets at the meaning of a word, not just what it should look like.
- ▶ You can pack lots of information within the transcription and then later use it to search texts.
- ▶ Depending on encoding, texts have different meanings.

Example:

I saw Washington, it was fabulous!

I saw <place reg="Washington, DC">Washington</place>, and it was fabulous!

I saw <person reg="Washington, George">Washington</person> and it was fabulous!

XML and Editing

Encoding is designed to capture the following:

- ▶ Metadata about the document
- ▶ The structure of the document
- ▶ The content of the document
- ▶ The function of words in the document
- ▶ The editorial process of creating the document
- ▶ The context of the document.

Examples:

```
<p>He said to meet us at <unclear  
sameAs="Maltow">Malthoo</unclear>  
Street.</p>
```

```
<gap extent="one page"  
reason="missing">
```

```
<placeName  
type="house">Baymeath</placeName>,  
<placeName type="city">Bar  
Harbor</placeName> <placeName  
type="state">Maine</placeName>
```

XML and Editing

Encoding can also make your transcriptions dance.

- ▶ You can let readers choose to see the original version of the text or the regularized.
- ▶ You can have identifications for words appear when you hover over them.
- ▶ You can provide translations for words in the encoding.
- ▶ You can embed several draft versions of a document in one transcription that can allow readers to toggle between versions.

Examples:

```
<choice>  
  <orig>favour</orig>  
  <reg>favor</reg>  
</choice>
```

Using XML

How do you determine how much to encode?

- ▶ Don't just do it because there is a tag or attribute for it.
- ▶ If it won't be used to drive your presentation of the document, searches, or other research function, consider leaving it out.
- ▶ The more complex your tagging, the longer it will take to do.
- ▶ Use templates to start documents off with most of the encoding in place and valid.

Sanger Project's use of XML

- ▶ We used the Text Encoding Initiative Guidelines (tei.org)
- ▶ We decided to encode mostly for content, not appearance.
 - ▶ We wanted subject access, when people, places, titles, or organizations were mentioned in the text.
 - ▶ We wanted to be able to search by title, date, type of document (speech, articles, etc.), and in what journal it was published.
- ▶ We decided to correct misspellings to make the searches more accurate. We could have encoded two versions, but chose not to.
 - ▶ Rationale was that many documents were published and errors were often not Sanger's.

Search screens

Advanced Search

Title:	(search through all of the document titles)	
Full Text:	(search through the body of the documents)	

Date: *(search on one date or from a range of dates)*

On/From:	Month	▼	Through:	Month	▼
	Date	▼		Date	▼
	Year	▼		Year	▼

Document Type 1:	(find documents of particular types)	▼
Document Type 2:	(find documents of particular types)	▼
Journal Title:	(find documents from a particular journal)	▼

Subject Index:	-- descriptions of -- MS on abortion -- birth control and -- frequency of -- health risks	⌵
----------------	--	---

Mentions:

Place:	(find documents that mention a particular place)	▼
Person:	(find documents that mentions a particular person)	▼
Organization:	(find documents that mention a particular organization)	▼
Title:	(find documents that mention the title of a work)	▼

Search!Reset

Challenges of XML

- ▶ Programming knowledge is helpful
- ▶ Publication platforms are challenging
- ▶ Harder to see what your documents look like while you are working on them.
- ▶ Changing guidelines. TEI P4, P5, and soon P6.
- ▶ In many ways it is overkill if you don't use the higher-end features.

MRS. SANGER HERE TO SPREAD GOSPEL OF BIRTH CONTROL

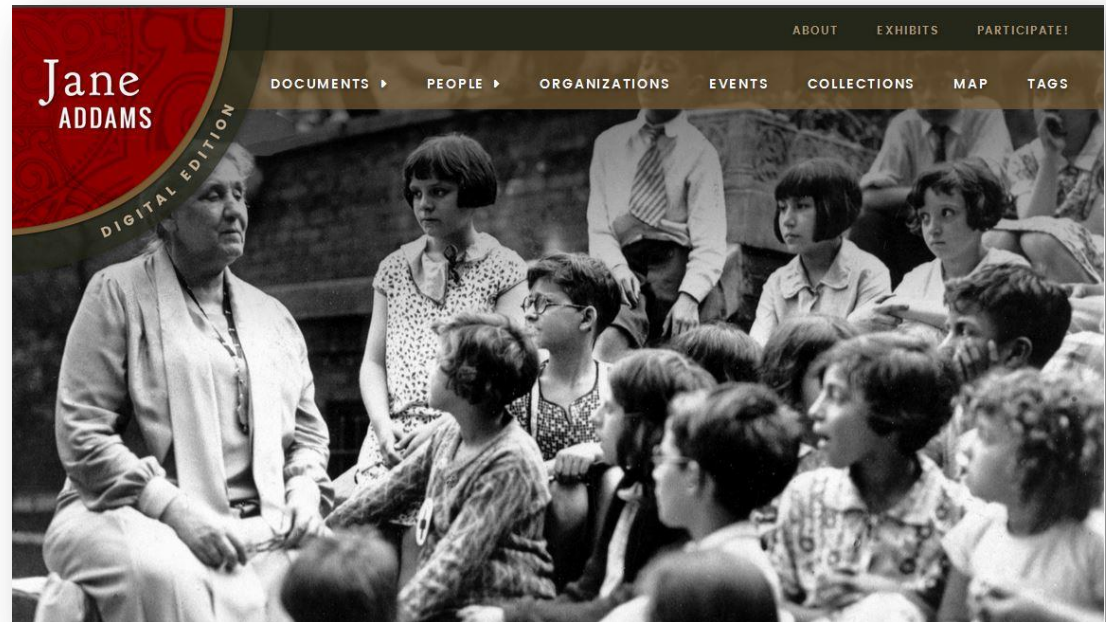
Mrs. Margaret Sanger of New York is in St. Louis to spread the knowledge of birth control. She will deliver a lecture and take steps toward the organization of a birth control clinic for the instruction of women. Mrs. Sanger, was the editor of the *Woman Rebel*, which was discontinued after seven of the nine issues had been confiscated by the postal authorities, and is the author of pamphlets on family limitation. A Federal indictment charging her with improper use of the mails was dismissed in February on the eve of her trial in New York. The reason given was that it was desired to avoid making a martyr of her. Mrs. Sanger, who arrived at the Jefferson Hotel this morning, said to a *Post-Dispatch* reporter that she had no wish to be a martyr but she wanted the case to go to trial, desiring acquittal at the hands of a jury and confident that she would have been acquitted.

Making Tour of Country.

She is touring the country in advocacy of birth control. She came here from Indianapolis, where she spoke at the national charity conference and where such interest was aroused, she says, that birth control teas and birth control luncheons are crowding out other society functions. Not that her message is to society women, she explained, except as she desires to interest them in the propaganda, which is intended to benefit the women of the laboring classes. "I have no message for wealthy women," she said. "They have been taken care of. The woman of Fifth Avenue and Riverside Drive controls her family. It is the woman of the poorer classes who needs enlightenment." "I was prompted by my experience as a nurse and as a mother to take up this propaganda. In my work as a nurse on New York's East Side I saw that there was overbreeding of the least desirable elements, which was not making for a racial improvement.

Jane Addams Papers

- ▶ Began in 2015
- ▶ Includes both images and transcriptions of about 22,000 letters, and writings.
- ▶ Uses short identifications of people, organizations and events to provide context.
- ▶ Uses maps and tags to aid searches
- ▶ Uses Omeka, a collections management platform that employs Dublin Core metadata.



Goals for digitization

- ▶ Digitizing portions of microfilm
- ▶ Transcription to make the documents more accessible.
- ▶ Building encyclopedia-like references to Addams's world.
- ▶ Use of standard metadata scheme allows us to export the data as XML for sharing and reuse.
- ▶ Welcome collaboration through crowdsourcing and exhibits.
- ▶ Taking an “off the shelf” platform and customizing it for the needs of scholarly editors.

TAGS



Omeka

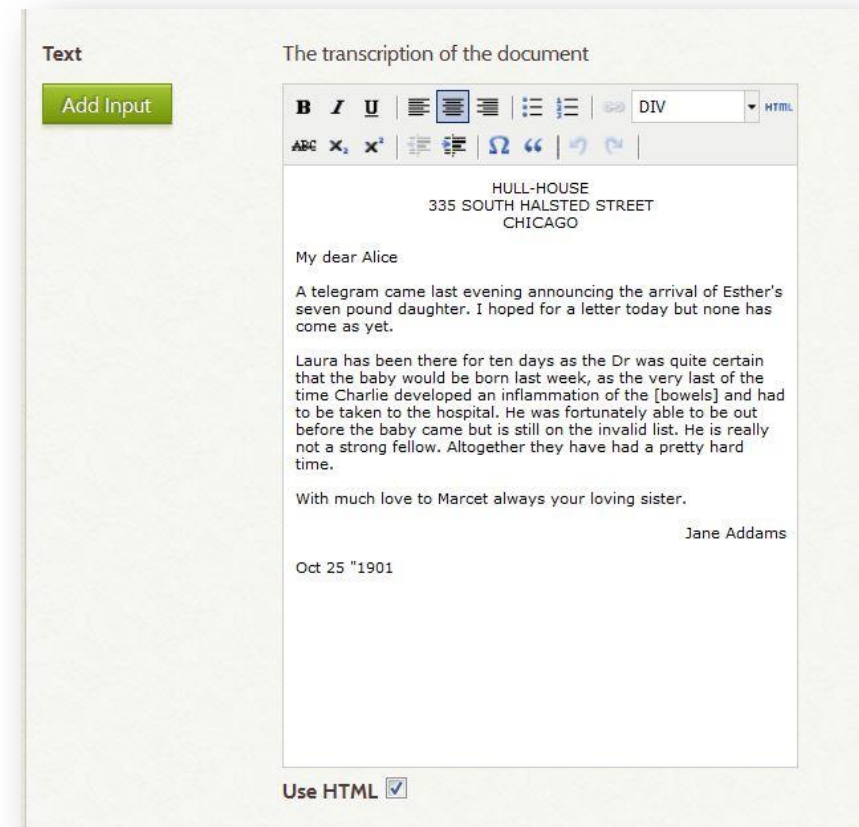
- ▶ Open-source, produced by George Mason University's Roy Rosenzweig Center for History and New Media
- ▶ Relies on relational databases to describe and link records
- ▶ Has an extremely easy web publication process (click a button)
- ▶ Built in searches
- ▶ Easy data entry
- ▶ Plugins allow customization of your site
- ▶ Omeka.org vs. Omeka.net vs. Omeka-S

The screenshot shows the Omeka-S administration interface for the 'Jane Addams Digital Edition'. The top navigation bar includes links for Plugins, Appearance, Users, Settings, and a welcome message for 'Cathy Moran Hajo'. A sidebar on the left lists various management tools like Dashboard, Items, Collections, and Reports. The main area displays a 'Browse Items' table with 6578 total items. The table has columns for Title, Creator, Type, and Date Added. Three items are visible in the list, each with a thumbnail, title, creator, type, date, and status indicators like 'Monitor' and 'Metadata Status'.

Title	Creator	Type	Date Added
Jane Addams to Sarah Alice Addams Haldeman, February 2, 1914 (Private) <i>Dublin Core</i> Date: 1914-02-02 Monitor Metadata Status: Incomplete Transcription Status: Pending	Addams, Jane	Text	Nov 15, 2016
A. Woods to Jane Addams, January 31, 1914 (Private) <i>Dublin Core</i> Date: 1914-01-31 Monitor Metadata Status: Complete Transcription Status: Pending	Woods, A	Text	Nov 15, 2016
Bruere, Henry, (?-?) (Private) <i>Dublin Core</i> Metadata Status: Incomplete		Person	Nov 15, 2016

Omeka

- ▶ Much simpler transcription capability
- ▶ Limits on what you can do with web presentation
- ▶ HTML produces a WYSIWYG transcription that you can easily see both in this admin view and by pressing a button, with the public theme applied.
- ▶ Works better combined with images.



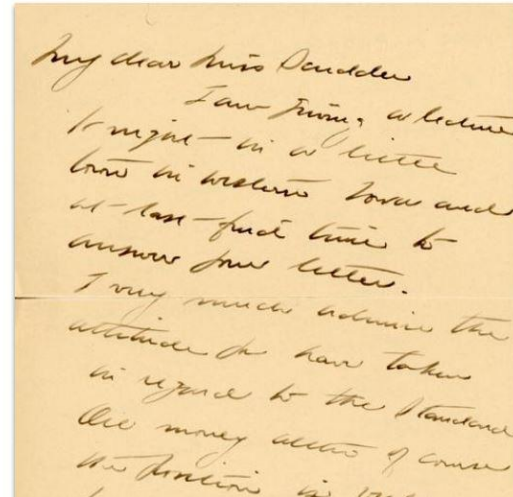
Transcription

- ▶ Enables searches
- ▶ Broadens accessibility
- ▶ Can be less complex because images are provided alongside the transcription
- ▶ Use some symbols to indicate inserted text, brackets to indicate where we have changed spellings or guessed at word.

JANE ADDAMS TO VIDA SCUDDER, APRIL 25, 1900

TRANSCRIPT

IMAGE



JANE ADDAMS TO VIDA SCUDDER, APRIL 25, 1900

TRANSCRIPT

IMAGE

My dear Miss Scudder

I am giving a lecture tonight in a little town in western Iowa and at last find time to answer your letter.

I very much admire the attitude you have taken in regard to the Standard Oil money [although] of course the position is vulnerable for many points. It really reduces itself to this--that if none of us formulate our opinions and compunctions act upon them, there [page 2] will be absolutely no chance of getting our and we must face the inconsistency of acting individually upon a social matter, until our groups are formed. Personally I am <not> so clear as you are in regard to this particular money--the University of Chicago presents new perplexities constantly--but I am thankful for your view and the statement of it.

CRE

ADDA

DAT

1900

DES

Adda

Welle

Oil, ar

settle

SUB

ADDA


SETTL

SETTL

Identifications

- ▶ We have already found over 2,500 individual people who wrote to Addams, received a letter from her, or was mentioned in a document.
- ▶ Aside from biographical information, we also link all mentions of a person to the biographical page.
- ▶ These serve as annotations for the digital edition.
 - ▶ References between documents are tracked (drafts, enclosures, replies)
 - ▶ Relationships between events, organizations and people are tracked (member of, participated in)

SCUDDER, VIDA DUTTON (1861-1954)



IMAGE

BIRTH DATE
1861-12-15

DEATH DATE
1954-10-09

BIRTHPLACE
MADURAI, INDIA

BIOGRAPHICAL TEXT
In her early life, Scudder and Clara French were the first two women admitted into Oxford's Un the Episcopate and into the general calendar on October 10.

Item Relations

RELATION FROM THIS ITEM

- Organization
 - Denison House (Member of)

RELATIONS TO THIS ITEM

- Text
 - Helena S. Dudley to Jane Addams, April 30, 1902 [Mentions]
 - Vida D. Scudder to Jane Addams, June 13, 1902 [Written by]
 - A. H. Nelson to Jane Addams, January 11, 1907 [Mentions]
 - Helena S. Dudley to Jane Addams, June 24, 1901 [Mentions]
 - Vida Dutton Scudder to Jane Addams, April 27, 1902 [Written by]
 - Louis F. Post to Jane Addams, October 1, 1907 [Mentions]
 - M. Carey Thomas to Jane Addams, May 7, 1908 [Mentions]
 - Marie Rollin Tuttle to Jane Addams, August 4, [1904] [Mentions]
 - Vida D. Scudder to Jane Addams, November 13, [1909] [Written by]
 - Vida D. Scudder to Jane Addams, July 19, 1912 [Written by]
 - Florence Converse to Jane Addams, February 22, 1910 [Mentions]
 - Jane Addams to Mary Rozet Smith, April 6, 1902 [Mentions]
 - Jane Addams to Vida Scudder, April 25, 1900 [Received by]

Scudder also co-founded the third settlement house in the United States, Denison House of Boston, with Helena Dudley and Emily Greene Balch in 1890. She served as its primary administrator from 1893-1913.

OCCUPATION
AUTHOR
PROFESSOR
SOCIAL REFORM ACTIVIST
RELIGIOUS ACTIVIST

COLLECTION
PEOPLE

TAGS
RELIGION
SOCIAL WELFARE
WOMEN
WRITING

Plugins

- ▶ Omeka has a range of plugins
 - ▶ Geolocation – creates a map
 - ▶ Item Relations – builds links between items that you design
 - ▶ Curator Monitor – allows control over the process of publishing materials.

Map

List of locations

All	Latitude	Longitude	Zoom Level	Map Type
<input type="radio"/>	41.8716782	-87.6474517	3	Roadmap
<input type="radio"/>	800 South Halsted S	Written from		


Add a new location

Find by address or point to a location

Address to find

Find

Add



Item Relations

Here you can relate this item to another item and delete existing of the relations, see the [Browse Vocabularies](#) page. Invalid item I

Subject	Relation	Object	
This Item	Written by	Addams, Jane (1860-1935)	<input type="checkbox"/>
This Item	Received by	Haldeman, Sarah Alice Addams (1853-1915)	<input type="checkbox"/>
This Item	Mentions	Hulbert, Esther Margaret Linn (1880-1955)	<input type="checkbox"/>
This Item	Mentions	Hulbert, Charles E. (1880-1939)	<input type="checkbox"/>
This Item	Mentions	Haldeman-Julius, Anna Marcet (1887-1941)	<input type="checkbox"/>
This Item	Mentions	Addams, Laura Shoemaker (1856-1937)	<input type="checkbox"/>

New Relations

Add a Relation

Monitor

Metadata used to manage various status about items.

Metadata Status Set by
Cathy Moran Hajo (cathymoranhajo) on 2016-04-15 08:54:41
Main status of metadata of the record, set by the staff member that created the document metadata/record.
This is used for documents and identifications.
Published

Transcription Status Set by
Cathy Moran Hajo (cathymoranhajo) on 2016-04-15 09:03:12
Status of the transcription, set by the staff member that did the initial transcription.
This is for documents only.
Published

Transcription Difficulty
Set by Tori Sciancalepore (vsclanca) on 2016-03-07 10:04:41
Information on the difficulty of the transcription according to the document.
For documents only. Easy is typed, Medium is poor quality typed or easy handwriting, and Difficult is most everything else.
Difficult

Subjects and Tags

- ▶ You can build a list of subjects and tags to chunk items into different groupings.
- ▶ We have some overlap between the two, but have tried with tags to use larger categories that can cover both the documents and the identifications.
- ▶ We have broken down the subjects relating to Jane Addams and Hull-House into narrower terms, because there are so many of them. This enables a book-like search.
- ▶ We also used occupation as a way to browse individuals.

A

ADDAMS, JANE, AND AFRICAN-AMERICANS (26)

ADDAMS, JANE, AND CHILD LABOR (46)

ADDAMS, JANE, AND FAMILY (195)

ADDAMS, JANE, AND HULL-HOUSE (24)

ADDAMS, JANE, AND IMMIGRANTS (33)

ADDAMS, JANE, AND LABOR MOVEMENT (34)

ADDAMS, JANE, AND PROGRESSIVE PARTY (102)

ADDAMS, JANE, AND PUBLIC HEALTH (4)

ADDAMS, JANE, AND SETTLEMENT MOVEMENT (26)

ADDAMS, JANE, AND THE GOVERNMENT (2)

ADDAMS, JANE, AND WOMAN SUFFRAGE (41)

ADDAMS, JANE, AUTOBIOGRAPHICAL VIEWS (10)

H

HEALTH WORKER (3)

HISTORIAN (34)

HOMEMAKER (51)

HOSPITAL ADMINISTRATOR (2)

I

ILLUSTRATOR (1)

INDUSTRIALIST (5)

INSPECTOR (12)

INSURANCE SALES (9)

INVESTIGATOR (3)

INVESTOR (1)

J

JOURNALIST (95)

JUDGE (44)

Future Expansions

- ▶ Crowdsourcing transcriptions
- ▶ Crowdsourcing rating items
- ▶ Digital exhibits drawn from the edition
- ▶ Text mining and topic modeling plugins
- ▶ Who knows?



Editing and Digital Humanities

- ▶ Audience broadens and we need to adapt to their needs.
- ▶ Platforms change over time and you need to choose them carefully and with a sure sense of your materials.
- ▶ Must be ready to pivot and include new tools as they arrive and become popular.

