

# DLF Aquifer: The Final Story

Katherine Kott, Susan Harum, Kat  
Hagedorn, Tom Habing  
DLF Spring Forum, Raleigh, NC  
May 5, 2009

# Aquifer accomplishments

- Research and best practices
  - Metadata guidelines
  - Tools inventory
  - Metadata crosswalks
  - Digital library assessment
  - Interoperability
  - Collaboration
- American Social History Online (funded project)
  - Open source code
  - Web site and services for scholars

# Best practices for collaboration

- Define and scope project/insure resource match
- Create a framework
- Define participation and governance
- Organize volunteer effort (skills)
- Be aware of participant priorities
- Select workable tools
- Celebrate success!

# The DATA

How we worked with the metadata  
you provided

Kat Hagedorn – [khage@umich.edu](mailto:khage@umich.edu)

# Many institutions contributed

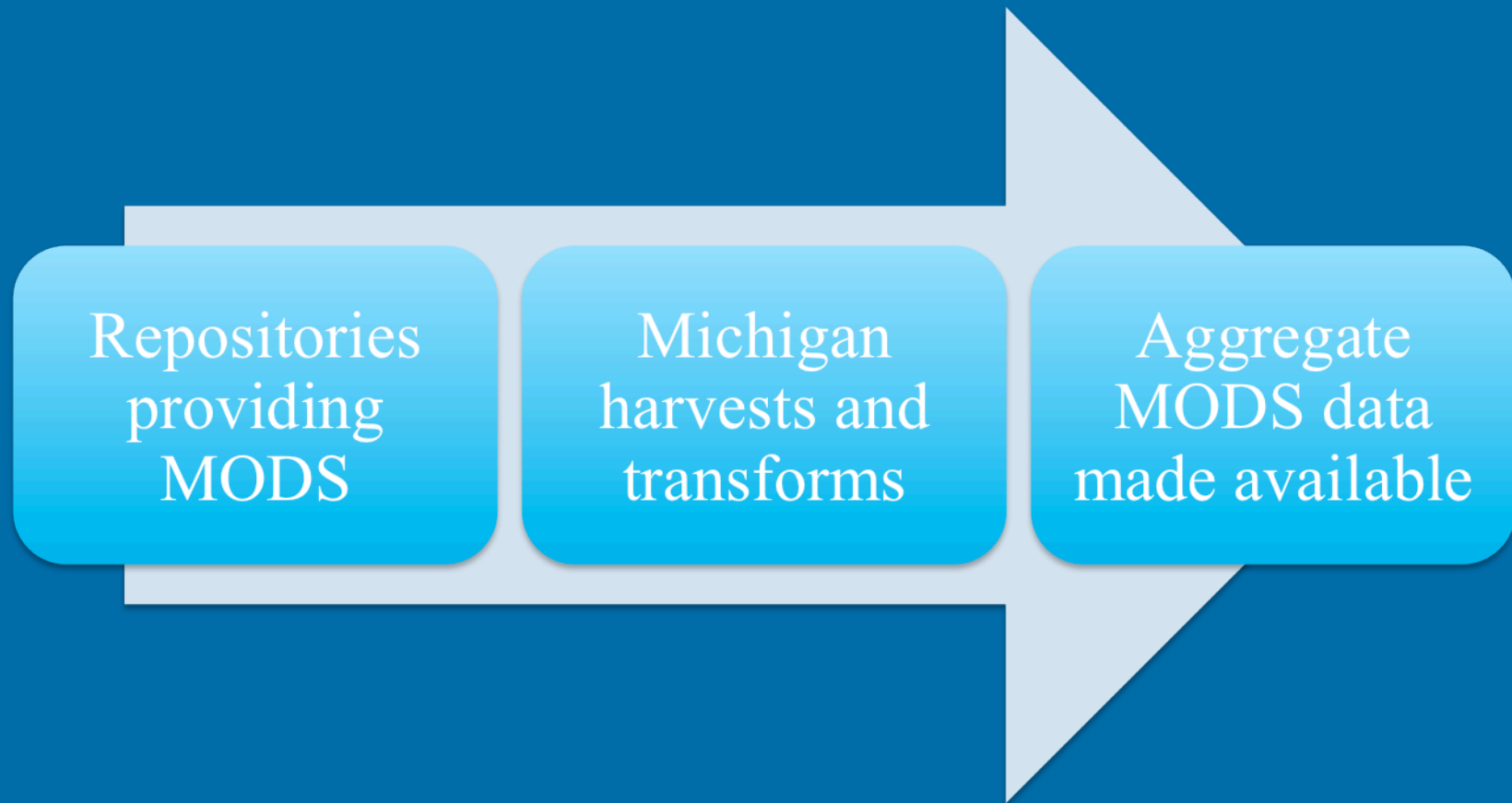
- California Digital Library
- Columbia University
- Cornell University
- Emory University
- Harvard University
- Indiana University
- Library of Congress
- New York University
- Northwestern University
- PBS / Florida Center for Library Automation
- Stanford University
- University of Chicago
- University of Illinois at Urbana-Champaign
- University of Michigan
- University of Minnesota
- University of Pennsylvania
- University of Southern California
- University of Tennessee at Knoxville

**THANK YOU!**

# The numbers

- 18 institutions
- 175 repositories (OAI sets)
- 514,708 MODS records
- 253,631 records describing images
- 175,990 records describing texts
  - similar to proportion of descriptive records in OAIster (majority of images)

# The big arrow is OAI



# Some things we changed




- Michigan performed some “added value” of the harvested records:
  - links to existing thumbnails, if available
  - an appropriate provenance container
    - e.g., `<originDescription altered="true" harvestDate="2008-07-22T:16:14:44">`
  - inclusion of data repository name
    - e.g., University of Michigan Library Repository:  
Making of America



# Hand-off

- Then it was Tom's— he harvested the aggregate MODS data and made it available at [www.dlfaquifer.org](http://www.dlfaquifer.org)
- But before that could happen...





# What metadata librarians live for

- Decisions on search, sort, display of fields
- Most problematic:
  - dates  keyDate, dateIssued, copyrightDate, dateCreated, ...
  - URL  mods/location/url vs. mods/identifier
  - subject  indexes for geographic and topical subjects (e.g., subject vs. genre)

# All ducks in a row

www.dlfaquifer.org



- Title** [The salamander safe. A millerite preparing for the 23rd of April](#)  
- Creator** Sinclair, Thomas S., ca. 1805-1881
- Date Issued** 1843
- Collection** Library of Congress Digitized Historical Collections: Cartoon Prints, American, 1766-1876
- Abstract** A playful caricature of a Millerite, an adherent of the Adventist preacher William Miller who predicted the Second Coming of Christ in 1843. The Millerites were known for their extreme beliefs and practices, including the use of a "Fire Proof Chest" stocked with a ham, a fan (hanging on the door of the safe), cheese, brandy, and other provisions. The "salamander safe," probably a trade name of the period, is named for the salamander, a creature believed to be immune to fire (without harm).
- Subject** [Miller, William; 1782-1849](#)  
[Millerite movement](#)
- Publication** monographic
- Genre** graphic; Caricatures.; Lithographs.
- Identifier** LC-USZ62-23784 DLC, hdl:loc.pnp/cph.3a24747
- Language** English
- Location** Library of Congress Prints and Photographs Division Washington, D.C. 20540 USA
- Note** Title appears as it is written on the item.  
[Printed by Thomas Sinclair, Philadelphia]  
Entered . . . 1843 by T[homas] Sinclair . . . Pa.  
The Library's impression of the print was deposited for copyright prior to the expected day of  
Published in: American political prints, 1766-1876 / Bernard F. Reilly. Boston : G.K. Hall, 199
- Physical Description** graphic; 1 print : lithograph on wove paper ; 26.6 x 29.4 cm. (image)
- Classification** PC/US - 1843.S613, no. 1
- Resource Type** still image
- URL** <http://hdl.loc.gov/loc.pnp/cph.3a24747>  
- Related Item** Weitenkampf, p. 73  
Murrell, p. 159

# MODS Levels of Adoption

- Created by the Metadata Working Group
- Each level indicated to data providers what could be done with their data
- Differing levels of conformance
  - in general Levels 1 and 2 were met more often than not, Levels 3-5 were more spotty
  - specifically– URLs, titles and type of resource were included by most repositories
  - keyDate used by only 14% of the repositories

# Over to Tom

- To test conformance to the Levels of Adoption, used samples of each institution's records in the MODS and Asset Action Explorer...

# Contact

- Thanks!
- Kat Hagedorn, [khage@umich.edu](mailto:khage@umich.edu)
- Digital Library Production Service,  
University of Michigan Libraries



# MODS and Asset Action Explorer

Check your MODS records against  
the DLF Aquifer MODS Guidelines  
Levels of Adoption

<http://ratri.grainger.uiuc.edu/aax/>

Tom Habing -- [thabing@uiuc.edu](mailto:thabing@uiuc.edu)


# Asset Actions

(not the primary focus of this presentation)

☒ View Asset Actions (2 records)

**Image Asset Action:** 19568FordFairlane [thabing] [11/6/2008 10:22:42 PM] [View](#) [Delete](#)


1958 Ford Fairlane Adieu



[View](#) [Collect](#)

**Image Asset Action:** BullRingJuarez [thabing] [11/13/2008 10:45:07 PM] [View](#) [Delete](#)

Bull Ring Juarez, Chihuahua



[View](#) [Collect](#)




Pages: 1



# MODS Levels of Adoption

<http://wiki.dlib.indiana.edu/confluence/x/q24>

1. Allows users to cite the resource
2. Allows users to perform basic searches and filtering
3. Allows users to browse and group search results
4. Allows users to perform more precise searches
5. Allows users to effectively evaluate resources

 REQ	"REQUIRED" designates a requirement of the group.
 REQIFAPP	"REQUIRED IF APPLICABLE" designates a requirement that is an absolute requirement if the condition is applicable to the resource.
 REC	"RECOMMENDED" designates a recommendation that an implementer may ignore without having weighed the implications.
 RECIFAPP	"RECOMMENDED IF APPLICABLE" designates a recommendation that is applicable if the condition is described and an implementer has fully weighed the implications.
 NOTREC	"NOT RECOMMENDED" designates a recommendation that an implementer may use, but it is discouraged after weighing the implications.
 OPT	"OPTIONAL" designates a recommendation that an implementer may use.

# Schematron

```
<sch:pattern id="p1" >
```

```
  <sch:title>Level 1</sch:title>
```

```
  <sch:p>Allows users to cite the resource.</sch:p>
```

```
  <sch:rule id="p1-mods" context="mods:mods">
```

```
    <sch:assert id="p1-mods-location" role="REQ"
```

```
      test="count(mods:location/mods:url[string-length(normalize-  
space(.))>0])>0">
```


One or more location elements with a url subelement is required.

```
  </sch:assert>
```

```
  </sch:rule>
```

```
</sch:pattern>
```

# Get an account

 ILLINOIS

MODS and Asset Action Explorer

[Login](#)

**Sign Up for Your New Account**

Note: Password must be at least seven characters and include at least one numeric and one special character, such as !@\$%+.

User Name:


Password:

Confirm Password:

E-mail:

[About This Service](#) [MODS Guidelines](#) [Asset Action Wiki](#) [Collectus Help](#)

# Upload some MODS records

 ILLINOIS MODS and Asset Action Explorer

thabing [Home](#) [Change Password](#) [Logout](#)

[My Uploaded Records](#) [All Uploaded Records](#) [All Uploaded MODS Collections](#)

MODS or Asset Action File    Name  (Only applicable to modsCollection uploads)


☐ Autogenerate a MODS recordIdentifier if needed.

XML Schema validation is against version 3.3 of the MODS XML schema; you may get an error if your MODS file has a version attribute different than this. You can safely ignore this error.

[About This Service](#) [MODS Guidelines](#) [Asset Action Wiki](#) [Collectus Help](#)

Will automatically be validated against MODS XML Schema

# Select a record to check

	ILLINOIS	MODS and Asset Action Explorer	thabing	<a href="#">Home</a>	<a href="#">Change Password</a>	<a href="#">Logout</a>
<a href="#">Upload New Record</a> <a href="#">My Uploaded Records</a> <a href="#">All Uploaded Records</a>						
<b>MODS Record:</b> dlc.12023971 <a href="#">Check Aquifer Levels of Adoption</a> <a href="#">Validate Against XML Schema</a> <a href="#">MODS XML</a>						
Uploaded by thabing on 10/31/2008 12:14:46 PM						
<b>Title Information</b>						
Title		Dud leaves home				
<b>Title Information:Alternative (displayLabel="Variant title in Library of Congress video collection:")</b>						
Title		Us fellers: "Dud leaves home"				
<b>Title Information:Alternative (displayLabel="Copyright title:")</b>						
Title		Goldwyn-Bray Pictographs, no. 7009				
<b>Name:Personal</b>						
Name Part		Carlson, Wallace				
Name Part:Date		d. 1967				
Role						
Role Term:Text		animation,				
Role						

# Check the record

## Level 1

Requirements for MODS Level of Adoption 1: Minimum for participation: Allows users to cite the resource.



**[REQ] count(mods:location/mods:url[string-length(normalize-space(.))>0])>0**

**p1-mods-location:** One or more location elements with a url subelement is required.

*Location: /\*[local-name()='mods' and namespace-uri()='http://www.loc.gov/mods/v3']*



**[REQIFAPP] count(mods:accessCondition[string-length(normalize-space(.))>0])>0**

**p1-mods-accessCondition:** The accessCondition element is required if anything other than completely unrestricted.

*Location: /\*[local-name()='mods' and namespace-uri()='http://www.loc.gov/mods/v3']*

## Level 2

# Select a whole collection to check

**I**LLINOIS MODS and Asset Action Explorer thabing [Home](#) [Change Password](#) [Logout](#)

[Upload New Record](#) [My Uploaded Records](#) [All Uploaded Records](#) [All Uploaded MODS Collections](#)

**MODS Collection DAG (146 records)** [Check Aquifer Levels of Adoption](#)

**MODS Record:** 2007675782 [thabing] [1/10/2009 10:02:26 AM] [View](#) [Delete](#)

[Stephen Arnold Douglas, head-and-shoulders portrait, slightly to left]

**MODS Record:** 2004664182 [thabing] [1/10/2009 10:02:27 AM] [View](#) [Delete](#)

[James J. Mapes, head-and-shoulders portrait, facing left]

**MODS Record:** 2004664011 [thabing] [1/10/2009 10:02:28 AM] [View](#) [Delete](#)

[John Alexander Kennedy, half-length portrait, three-quarters to left, eyes front]

**MODS Record:** 2004663911 [thabing] [1/10/2009 10:02:28 AM] [View](#) [Delete](#)

[Jacob Collamer, half-length portrait, three-quarters to left]

**MODS Record:** 2004663870 [thabing] [1/10/2009 10:02:29 AM] [View](#) [Delete](#)

[Bishop Frederic Baraga, three-quarter length portrait, facing three-quarters to right, seated, in clerical robes, holding his Dictionary of the Otchipwe Language]

# Check them all and create a report

ILLINOIS MODS and Asset Action Explorer thabing [Home](#) [Change Password](#) [Logout](#)

[Upload New Record](#) [My Uploaded Records](#) [All Uploaded Records](#) [All Uploaded MODS Collections](#)

**Aquifer Levels of Adoption for MODS Collection 52 (146 records)**  
*Hover your mouse over the Assert/Report IDs and Roles for more information.*

**Pass/Fail Record Counts By Rule** [Download CSV](#)  
(passed plus failed plus not checked equals the number of records)




Assert/Report ID	Role	Records Passed	Records Failed	Not Checked/Applicable
p1-mods-accessCondition		146	0	0
p1-mods-location		146	0	0
p1-mods-name		146	0	0
p1-mods-originInfo		146	0	0
p1-mods-titleInfo		146	0	0
p2-mods-language		146	0	0
p2-mods-physicalDescription		146	0	0
p2-mods-physicalDescription-internetMediaType		0	146	0
p2-mods-subject		96	50	0

Download as a CSV if desired



# ...or two reports

Pass/Fail Rules By Record [Download CSV](#)

Record ID	p1-mods-accessCondition 	p1-mods-location 	p1-mods-name 	p1-mods-
<a href="#">2004663863</a>	Passed	Passed	Passed	F
<a href="#">2004663864</a>	Passed	Passed	Passed	F
<a href="#">2004663866</a>	Passed	Passed	Passed	F
<a href="#">2004663868</a>	Passed	Passed	Passed	F
<a href="#">2004663870</a>	Passed	Passed	Passed	F
<a href="#">2004663873</a>	Passed	Passed	Passed	F
<a href="#">2004663878</a>	Passed	Passed	Passed	F
<a href="#">2004663881</a>	Passed	Passed	Passed	F
<a href="#">2004663882</a>	Passed	Passed	Passed	F

...

...

<a href="#">2007675782</a>	Passed	Passed	Passed	F
RECORDS FAILED COUNT	0	0	0	

...

Download as a CSV if desired

# Rinse and repeat

- To protect the innocent, delete your embarrassingly nonconformist records
- Fix all the problems
- Upload them again and check for errors
- Repeat until you are 100% conformant

# Celebrate

- By making your records OAI harvestable and having them added to the DLF ASHO Portal.
- Register with the UIUC OAI data provider registry:
  - <http://gita.grainger.uiuc.edu/registry/>
- and the DLF Collections Registry:
  - <http://dlf.grainger.uiuc.edu/dlfcollectionsregistry/>

# Thanks

Tom Habing – [thabing@uiuc.edu](mailto:thabing@uiuc.edu)



University of Illinois at Urbana-Champaign  
Grainger Engineering Library Information Center

# Assessment Activities Final Report

Susan Harum

Business Analyst

American Social History Online

# Assessment Activities 2008-2009

- Interviews
- Focus Groups
- Surveys
- Semester long classroom study
- Google Analytics data

# Expected Outcomes

- *Improving Access for Scholars*
- *Stimulating New Research Questions*
- *Supporting Interdisciplinary Study*
- *Supporting Cross-Regional Research*
- *Increasing Digital Collection Use*
- *Model Development.*

# Methodology- SWG

## Collaborative Work

- *ASHO* web site and Commercial Search — Katherine Skinner, Emory University, Chris Bourg, Stanford University, and Deb Holmes-Wong, University of Southern California
- *ASHO* web site and Metasearch — Susan Harum, University of Illinois at Urbana-Champaign
- *ASHO* web site and Zotero — Jason Phillips, New York University
- Google Analytics analysis of *ASHO* use — Susan Harum, University of Illinois at Urbana-Champaign



# ASHO Strengths

- Aggregated content across different archives that provide cross-collection searching
- The “vettedness” of the materials
- The number of visual resources (images)
- The metasearch feature
- The faceted browsing and searching, especially the facets associated with dates

# Feature Strengths

- Collection pages – good look and feel and useful for determining value of individual collections quickly
- Thumbnails on the results pages - used for determining the usefulness of the resource
- Metasearch feature - in particular the rare and important resources available through Google Book Search
- Browsing and searching facets

# Overall

Interviewees indicated that they would willingly use this resource now in their teaching and in their research so long as the **content body continued to increase** such that they would find more relevant resources over time.

Experienced researchers and graduate students both saw *American Social History Online* as a valuable tool for use in the classroom and as a set of services that could be used to introduce undergraduates to the wealth and range of primary material that is available

# ASHO Weaknesses

- The small size of the *American Social History Online* collection
- Interface issues

# Overall

- Interviewees indicated that they would willingly use this resource now in their teaching and in their research so long as the content body continued to increase such that they would find more relevant resources over time.

# Google Analytics

- What outreach and communication efforts help referral trends so that we can promote the site efficiently?
- What features are important to the users?
- Does the ASHO site help promote traffic to our partnering organization's (collection contributors) sites?
- Is it possible to see if the MODS levels of adoption help users find items?

# Traffic

- ASHO site has had a consistent increase in traffic, both overall and of its pageviews, including an increase in the length of time users are spending on the site.
- The number of increased pageviews and the increased time spent on the site means not just that traffic is up but that users are finding useful content.

# Tracking Referrals – what has worked

- Blogging by project members (on Aquifer project blog and individual project member blogs)
- Referrals by DLF site
- Entries on libsite.org, Wikipedia digital libraries page (for history)
- Project slides on slideshare.com
- Links from participating institution sites



# Outreach?

- Inconclusive, but:

Data shows a consistent improvement in traffic with a more dramatic increase at the end of October 2008

High number of users typing address in directly (equals referrals from other sites) could be an indicator of presence at historian and library conferences have helped communicate about the site

Anecdotal evidence from assessment and conferences demonstrates that increasingly the site is being found and used by historians, especially those needing visual and primary sources for teaching

Large jump in users at the end of October 2008 after concentration of outreach and assessment activities

## Features Most Used ( From Top Content, Entrance Paths and Map Overlay)

- Quick Search feature on home page
- Faceted browsing feature on the homepage, which allows the user to browse subjects, places and dates (by century)
- Collection pages (browse content and verify provenance)
- Facets for breaking down a search

# Features Not Widely Used

- Openid login
- Zotero
- Collectus
- Timeline
- Geomap

# Top Exit Pages

- Majority of visitors leave after investigating the home page, executing a single quick search, or browsing through the collections page.
- However, those that end up spending more time on the site go to specific items in the database.
- The Top Exit pages feature exits to unique items by the fifth entry on down. This means that the ASHO site helps to promote traffic to our collection contributors' site where the item is maintained.

# Metadata?

- Faceted browsing feature on the homepage most popular place for users to begin after quickly using the quick search feature (LOA level 3 - <subject> with specific subelements, i.e. <geographic> ("highly recommended" in Guidelines; LOA level 3). This was used to create the Places browse from the front page, and also the place, city and state filters to limit search results.
- Content Drilldown pageviews whose "search URL" contained "decade" or "year" comprised 13.9% of the total 35,029 pageviews.
- Pageviews containing "geographic" "city\_state" or "=state" comprised 14.5% of pageviews. (Note that some pageviews utilized both temporal and geographic facets or facet tags).

# Conclusion

- Assessment shows that *American Social History Online* collections and services are perceived as useful, especially in introducing undergraduate students to the wealth of primary materials available for study.
- Continuing to grow and develop *American Social History Online* collections services offers interesting challenges and opportunities in keeping both scholars and the digital library community interested in *American Social History Online*

# Where is Aquifer now?

- Products developed through American Social History Online being maintained by University of Illinois at Urbana-Champaign
  - Web site
  - Collections registry
- Standards and best practices
  - MODS metadata guidelines migration to LoC
  - Other documentation???

*Thank you for your interest*

- Questions or comments?
- For more information--[kkott@stanford.edu](mailto:kkott@stanford.edu)