

EAD Finding Aids and Search

San Jose State University

INFO 244 | Final Project

Greg Reeve

12/11/2015

Introduction

Archives and Special Collections

Finding aids = Archival Descriptions

Finding Aids

- Library's and archivists
- Researchers and patrons

Preserve, store, find, discover, identify, retrieve.

*Finding aids are the
"workhorse of archival practice"*

(Gilliland-Swetland, 2001, p. 200).

EAD

Encoded Archival Description

XML

Extensible Markup Language


```
<!--?xml version="1.0" encoding="UTF-8"?-->
<!-- ===== Enhanced: =====
  Transformed with schema2schema.06_mfer.xsl
Generated: Mon, 20 July 2015 12:04:19 PM EDT
URL: http://hdl.loc.gov/loc.gmd/eadgmd.gm015006
===== --><!--name="eadidNode" select="normalize-space(//ead:eadid)

<ead xmlns="urn:isbn:1-931666-22-9" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:sch
  <eadheader repositoryencoding="iso15511" relatedencoding="MARC21" countryencoding="iso3166-1"
    <eadid mainagencycode="US-DLC" countrycode="US" identifier="hdl:loc.gmd/eadgmd.gm015006" er
  <filedesc>
    <titlestmt>
      <titleproper encodinganalog="245$a" > [Waldo R. Tobler collection: the beginnings of
        computerized cartography and geographic information systems]. </titleproper>
      <subtitle>A Finding Aid to the Collection in the Library of Congress </subtitle>
      <author>Prepared by Ryan Moore</author>
    </titlestmt>
    <publicationstmt>
      <publisher>
        <extptr xmlns:xlink="http://www.w3.org/1999/xlink" xlink:type="simple"
          Division<lb> Library of Congress</lb></extptr></publisher>
      <address>
        <addressline>Washington, D.C.</addressline>
      </address>
      <date encodinganalog="260$c" normal="2015" era="ce" calendar="gregorian">2015</date>
    </publicationstmt>
    <notestmt>
```

collection : the beginnings of computerized geographic information systems]

 all words

Search

[Search All Findings](#)[Index Terms](#)[Using this Collection](#)[Search Results](#)[Print/Download](#)[Home](#) | [Biographical/Organizational Note](#) | [Scope and Contents](#) | [Arrangement](#)

Search and EAD

[Waldo R. Tobler collection : the beginnings of computerized geographic information systems]

1954-1978

G70.2 coll Tobler

Tobler, Waldo R.

42 folders ; 1 archival box (.5 linear ft.)

Collection material in English

[Geography & Map Division, Library of Congress, Washington, D.C.](#)

Materials from Waldo R. Tobler's time as an Assistant Professor to Professor at the University of Michigan from 1961 to 1977 and portions of his thesis when a student at the University of Washington, from 1954 to 1961. Items include manuscripts, some of which were published, manuscript maps, early cartographic computer programs, instructional materials. Tobler's early research, (see especially this collection, notes on map projection), and work for making cartographic data into computerized simulation, both of which served as precursors to geographic information systems ("GIS"). Many of these papers are out of print and hard to find in libraries. Many are out of print and hard to find in libraries.

Cite or bookmark this finding aid as: <http://hdl.loc.gov/loc.gmd/eadgmd.gm015006>

LC Online Catalog record for this collection: <http://lccn.loc.gov/2015587099>

Searching for EAD Finding Aids

Successful searching

- Search database/index size
- Search engine feature set
- Researcher skills

Search challenges

- Subject-based search
- Search term selection
- Archival jargon
- Access points: there's too many of them!

Custom EAD search interfaces

Search strategies and features

- Subject
- Name
- Date
- Geographic location
- Physical form or genre
- Top-down
- Bottom-up

Conclusion

Thanks!

References

Allison-Bunnell, J., Yakel, E., %26 Hauck, J. (2011). Researchers at work: Assessing needs for content and presentation of archival materials. *Journal of Archival Organization*, 9(2), 67-104. doi:10.1080/15332748.2011.598400

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. Retrieved from <https://pages.gseis.ucla.edu/faculty/bates/berrypicking.html>

Brown, G. %26 Harvey, K. (2007). Adding archival finding aids to the library catalogue: Simple crosswalk or data traffic jam?

Chapman, J. C. (2010). Observing users: An empirical analysis of user interaction with online finding aids. *Journal of Archival Organization*, 8(1), 4-30. doi:10.1080/15332748.2010.484361

Coats, L. R. (2004). Users of EAD finding aids: Who are they and are they satisfied? *Journal of Archival Organization*, 2(3), 25-39. doi:10.1300/j201v02n03_03

Cornish, A. %26 Merrill, A. (2010). Enabling integrated, online access to special collections. *Technical Services Quarterly*, 27, 151-163.

Cornish, A. (December 2004). Using a native XML database for encoded archival description search and retrieval.

Cox, E. %26 Czechyowski, L. (2007). Subject Access Points in the MARC Record and Archival Finding Aid: Enough or Too Many? *Journal of Archival Organization*, 5(4), 51-59.

Daniels, M. G. & Yakel, E. (2010). Seek and you may find: Successful search in online finding aid systems. *The American Archivist*, 73, 535-568. Retrieved from <http://www2.archivists.org/american-archivist>

Duff, W. & Stoyanova, P. (1998). Transforming the crazy quilt: Archival displays from a users' point of view. *Archivaria*, 24, 44-79. Retrieved from <http://journals.sfu.ca/archivar/index.php/archivaria/index>

Eidson, M. Y. (2002). Describing Anything That Walks: The Problem Behind the Problem of EAD. *Journal of Archival Organization*, 1(4), 5-28.

Feeney, K. (1999). Retrieval of Archival Finding Aids Using World-Wide-Web Search Engines. *The American Archivist*, 62(2), 206-228.

Gilliland-Swetland, A. J. (2001). Popularizing the finding aid. *Journal of Internet Cataloging*, 4(3-4), 199-225.
doi:10.1300/J141v04n03_12

Hensen, S. L. (2001). Archival cataloging and the internet: The implications and impact of EAD. *Journal of Internet Cataloging*, 4(3/4), 75-95.

Hostetter, C. J. (2004). Online finding aids: Are they practical? *Journal of Archival Organization*, 2(1-2), 117-145.
doi:10.1300/j201v02n01_09

Library of Congress. (2012, July 11). About EAD. Retrieved from <http://www.loc.gov/ead/eadabout.html>

- MacNeil, H. (2012). What finding aids do: archival description as rhetorical genre in traditional and web-based environments. *Archival Science*, 12, 485-500. doi: 10.1007/s10502-012-9175-4
- Prom, C. J. (2004). User interactions with electronic finding aids in a controlled setting. *The American Archivist*, 67(2), 234-268. Retrieved from <http://www.jstor.org/stable/40294278>
- Tibbo, H. R. & Meho, L. I. (2001). Finding finding aids on the world wide web. *The American Archivist*, 64(1), 61-77.
- Yakel, E. (2004). Encoded archival descriptions: Are finding aids boundary spanners or barriers for users? *Journal of Archival Organization*, 2(1), 63-77. doi:10.1300/J201v02n01_06
- Zemsky, R. & Wegner, G. (2007). Changing roles of academic and research libraries. Retrieved from <http://www.ala.org/acrl/issues/value/changingroles>
- Zhou, X. (2007). Examining search functions of ead finding aids web sites. *Journal of Archival Organization*, 4(3-4), 99-118. doi:10.1300/J201v04n03_06

Thanks again!

INFO244 Final Project
SJSU
Greg Reeve
12/11/2015

Script

Hi, my name is Greg. In this presentation I will introduce the EAD finding aid and explore search-related issues connected with EAD finding aids.

Introduction

Archives and special collections perform important roles within an academic library.

They work to collect, to catalog, to store, and to preserve rare, valuable, or culturally significant items.

These collections curated by archivists enables researchers and library patrons continued access to these archival materials for scholarly and personal research.

To help researchers and library patrons as well as ensure the long-term preservation of archival holdings and collections, archivists create detailed descriptions of items and collections in their archives.

These archival descriptions are referred to as finding aids.

Finding aids are multidimensional in nature:

They serve the library and archivists' needs to preserve and store archival materials while simultaneously serving the needs of researchers and patrons in finding collections and items for research.

The finding aid is used by archivists as well as patrons to know what is in an archive and where to find the materials in the archive's collection (Duff & Stoyanova, 1998, p. 44).

Finding aids are the "workhorse of archival practice" (Gilliland-Swetland, 2001, p. 200).

These archival descriptions have been encoded in many forms throughout time including clay tablets and printed paper.

The adoption of computer and internet technologies in the late twentieth century enabled the creation of new electronic encoding formats for finding aids.

The Encoded Archival Description (EAD) is a digital encoding format for finding aids in an online environment (The Library of Congress, 2012).

EAD is a relatively new encoding format.

Before EAD was created the library catalog system was used to describe archival materials and collections.

The creation of EAD began in 1993 from a project started at the University of California in Berkley.

Participants in this project wanted to create a new, non-proprietary encoding standard that is machine readable and better describes archival holdings beyond what library catalog formats provided.

EAD was based on the computer markup language SGML, which was a precursor to HTML.

The EAD format is currently defined using Extensible Markup Language or XML.

Here is an example of an EAD finding aid in its raw XML format.

This finding aid comes from the library of congress.

The XML format is made up of various structured tags that describe the content that the tags wrap.

As you can see the XML format is hierarchical in nature.

The data is presented in multiple levels.

(scroll through the XML to “Extent”)

(describe extent)

(scroll through to the end and show that the files and folders are then described in low-level detail)

This is a view of the same EAD finding aid from the previous slide displayed in the Library of Congress’ custom EAD display interface.

EAD allows variability in display style (The Library of Congress, 2012) while being well structured and easily parsed by machines and read by humans.

EAD and the World Wide Web extends the availability of the finding aid potentially reaching more patrons outside an archive's local community (Coats, 2004, p. 32).

By using an EAD on the World Wide Web patrons from across the globe can know what's in an archive without actually having to go to the archive in person.

The use and display of EAD finding aids is an emerging field of study with a high degree of interest in the academic community (Hostetter, 2004, p. 123).

Search and EAD

There are various issues related to search and EAD finding aids that are identified in the research literature.

These issues can be categorized into two main areas:

1. Searching for EAD finding aids
2. Custom EAD search interfaces

Searching the web for EAD

The EAD combined with the World Wide Web provides better searching particularly aligned with Bates' "berrypicking"

search technique (Gilliland-Swetland, 2001).

EAD can provide hyperlinks to related items or digital representations of archival materials (Gilliland-Swetland, 2001).

Successfully finding a relevant EAD is influenced by many factors including

- the size of the search database index
- The feature set of search engines
- The skills of the searcher

(Tibbo & Meho, 2001, p. 63)

Keyword searching is currently the best supported style of searching for an EAD but requires prior knowledge of what to find using the EAD.

Research shows that precise known-item searches returned better results than subject-based searches (Daniels and Yakel, 2010).

In this way search becomes a barrier to success when trying to find an EAD (Daniels and Yakel, 2010).

Difficulty choosing search terms and lack of familiarity with EAD jargon are obstacles to successful searching (Daniels and Yakel, 2010).

The challenges of finding an EAD on the World Wide Web has led some researchers to argue for better cataloging practices of finding aids so that patrons connect with an EAD through the library cataloging system rather than directly through a search engine (Cox & Czechowski, 2007; Brown & Harvey, 2007).

Another challenge lies in there being so many possible access points for identifying an EAD that it's hard to know what search engines should use for search and retrieval versus doing a full-text search against the entire content in the EAD.

This leads to high recall but low precision in online search results.

EAD search interfaces

Finding aids can be made available in their raw XML form but are usually transformed to a format that is more friendly for human viewing and browsing.

Users typically interact with EAD finding aids through a custom EAD interface.

In 2007 a content analysis study was performed focusing exclusively on search features within an EAD display interface.

Results from this study showed that search capabilities varied widely across finding aids systems.

Varying implementations of search can lead to confusion where patrons aren't sure what to expect when searching for finding aids (Zhou, 2007, p. 117).

Other studies have found that current EAD search tools are effective for archivists and computer experts but became barriers to successfully completing search tasks for novice users (Prom, 2004; Yakel, 2004).

One researcher recommends based on their study results that EAD interfaces should enable a broader set of access points (e.g. geographic names or names of people) to search against.

Aligned with the Bates (1989) model of search strategies EAD interfaces should include a range of features to help users access what they seek.

Some key strategies include:

- subject searching
- Name searching
- Date searching
- Geographic location searching
- Physical form or genre searching
- Top-down and bottom-up searching

Conclusion

The overall consensus points towards the need for more research related to the EAD search interface.

Technology changes rapidly. Many studies are possibly outdated and need to be reconsidered with today's fast moving technology environment.

The original vision for EAD included the creation of a "union database" that would be the one-stop-shop for finding EADs across multiple archives and special collections.

This vision has yet to be realized.

What exists currently are many disparate systems with their own standards for displaying and searching EAD finding aids.

Researcher's focus on the design of EAD interfaces has missed understanding user information seeking behaviors that could lead to improving search for EAD finding aids (Daniels and Yakel, 2010).

Hierarchical searching as well as comparing the effectiveness of keyword and subject searching are key areas identified by the literature as needing further study (Allison-Bunnell et al., 2011, p. 95; Chapman, 2010; Daniels and Yakel, 2010; Prom, 2004; Zhou, 2007).

Another future area of research includes exploring existing computer science research into search and XML documents.

Looking at research within the field of computer science about search of XML formatted documents may be helpful to find new perspectives that weren't considered previously.

Focusing on understanding effective ways to search and display the EAD finding aids will lead to more effectively meeting the needs of EAD users.