Metadata Standards: A Review of the EAD (Encoded Archival Description) Standard

James Gross, Drexel University
Assignment#2: Metadata Standards:

A Review of the EAD (Encoded Archival Description) Standard

James Gross

Drexel University, Info 653

Dr. Ding

10/13/2013
This paper will briefly review the EAD metadata standard. It is a standard for describing collections held by archives and library special collections. In this paper we will briefly discuss and review the focus and background, as well as the purpose and significance of this metadata schema. We will briefly discuss the functions that this metadata supports, whom uses it, and whether any challenges exist in its implementation.

**Focus:** The focus of this paper is the Encoded Archival Description (EAD) standard. This is an SGML/XML (Standard Generalized Markup/Language Extensible Markup Language) document type definition (DTD) used for archival finding aids (Yankel & Kim, 2005, p. 1427). The EAD standard is supported by the Society of American Archivists, and the Library of Congress (Florida Center, 2009).

Per the Library of Congress (LC) website, the EAD standard: “…stands for Encoded Archival Description, and is a non-proprietary de facto standard for the encoding of finding aids for use in an online networked environment” (Library 2013a). The EAD standard was designed to enable users to locate and make use of a collection’s finding aid.

The LC defines finding aids as “…inventories, indexes, or guides that are created by archival and manuscript repositories to provide information about specific collections. (Finding aids) …provide detailed description of the content and intellectual organization of collections of archival materials. EAD allows the standardization of …information in finding aids within repositories” (Library, 2013a).

The finding aid, an integral part of the EAD standard, is the main tool for archival description. “A finding aid is a tool created by records professionals to give information about the contents of archival collections. Finding aids provide detailed inventories to help researchers locate relevant materials” (Bohr, 2013, n.p.). As Thurman noted, the finding aid,” …allows for a much more detailed presentation of an archival entity’s context and hierarchical arrangement” (2005, p.185).

**Background:** Archival institutions, depending on their collection emphasis, typically house collections of corporations, individuals, or organizations. Collection emphasis and subject matter are often based on the mission of the archive. These materials may include unique or unpublished records of activities of organizations, families, or individuals. These records are valuable information resources that, prior to the EAD schema, may have been difficult to identify and locate on the internet.

**History** Development of the EAD standard began with a project initiated by the University of California, Berkeley Library, in 1993 (Library, 2013b). This project, also known as the “Berkley Project” was led by Daniel Pitti, to develop a nonproprietary standard for encoding finding aids (Thurmond, 2005, pp. 186-7).

The goal of the EAD Berkeley Project was to “investigate the desirability and feasibility of developing a nonproprietary encoding standard for machine readable finding aids such as inventories, registers, indexes, and other documents created by archives, libraries, museums, and manuscript repositories to support the use of their holdings” (Library, 2013b).

In the early 1990s, finding aids for archival materials existed mainly in paper copies held by the repositories of those materials (Thurman, 2005, p. 186). This made the locating of collections very difficult and time consuming. The main problem facing archivists was that there was an inability to access all the descriptive information available in finding aids.
Archivists realized that there was a need to provide greater access to finding aids as well as the descriptive information available within the finding aids. However, existing access, via the format of microfiche, was non-machine-readable and not searchable.

**Purpose:** The purpose of the EAD standard was to “enable the encoding of archival finding aids into records that would be machine readable, and fully searchable” (Thurmond, 2005, p. 184).

This encoding standard was needed to enable computer defined syntax to differentiate between information describing the provenance of an archival fond, from information about its scope and content, or the difference between a personal name and a corporate name. Such distinctions were critical for search, retrieval, and indexing (Thurmond, 2005, p. 186).

Existing html encoding is often not sufficient to enable the searching of finding aids. As Thurmond noted, “HTML is an SGML (Standard Generalized Markup Language) and is intended only for displaying hypertext on the Web. HTML cannot capture the underlying intellectual structure represented in a finding aid” (Thurmond, 2005, p. 186).

**Why developed:** EAD was developed as an international standard for archives and manuscript libraries to encode data describing corporate records and personal papers (Pitti, 1999, p. 1). Archival records are traditionally catalogued by an archive with the creation of a finding aid. The goal was to standardize the encoding of finding aids.

The Berkley Library project had the following goals for the EAD standard: 1) the ability to present extensive and interrelated descriptive information found in archival finding aids, 2) the ability to preserve the hierarchical relationships existing between levels of description, 3) the ability to represent descriptive information that is inherited by one hierarchical level from another, 4) the ability to move within a hierarchical informational structure, and 5) support for element-specific indexing and retrieval (Library, 2013b).

Due to the complexity of some archival collections, EAD was created to support hierarchical description. A central feature of EAD is the ability to support mapping to and from other data standards including MARC and Dublin Core (Combs, 2010, pg. 9). An example of encoded archival content is at upper right (Figure #1, upper right).

**Significance:** The concept behind standardization was: To make it possible to create “union access” and indexes to archival descriptions, regardless of where the repository is located (Cornish, 2004, p.181). The goal was to enable libraries and archives to share information about related but different, records and collections (Pitti, 1999, p. 2).

EAD represents a highly structured way to encode and mount digital finding aids on the www (Yakel & Kim, 2005, p. 1427; DCMI, 2013, n.p.).

**What does EAD do?** EAD tags describe the characteristics of archival materials. (Partial EAD tag structure, Figure #2, at right).

```
4. EAD Required Tags: EAD was deve
There are only a few required tags:
<ead
<eadheader
<eadid
<filedesc
<titlestmt
<titleproper
<archdesc level attribute
<dsc type attribute
```
EAD makes use of a tag structure that identifies the components of a document. Each component or part is identified, and noted through the encoding. Because EAD is an application of XML, EAD utilizes the concepts of tags, elements, and attributes for encoding text (University, 2013).

**How does EAD work?**

EAD is based on the multi-level description of finding aids and collections used by archival repositories. Finding aids may include a general description of all materials, a description of series or groups, and a description of files or items. EAD encoding is based on, and seeks to reflect, the content and structures of archival finding aids (University, 2013).

The image (Figure #3, at left) is an example of the front-page indexed content created from a EAD finding aid. Pictured is: “Inventory of the Department of Aging,” from the Online Archive of California. It can also be found here: [http://tinyurl.com/mq7366z](http://tinyurl.com/mq7366z)

EAD is a hierarchical schema in which elements are nested within one another. An example of an EAD encoded Finding Aid template (Figure #4, not shown): [http://tinyurl.com/nxp52s5](http://tinyurl.com/nxp52s5)

**Within the EAD schema, some elements can contain text directly, while other elements are intended to help structure the finding aid into sections and cannot take text directly but instead must contain elements inside them, these structural elements are called “wrapper elements” (Thurman, 2005, p. 188).**

The EAD schema has helped to standardize archival descriptive practices while increasing progress toward a collective online access to the finding aids of archival materials (Thurmond, 2005, p.184).

This standard supports archival collections and libraries containing special collections. The primary interface is via the encoded finding aid. The primary benefit of this metadata standard is that researchers, regardless of geographic location, will be able to learn about the existence of an archival collection, based on the EAD encoded finding aid. This will benefit anyone doing any special research on any person, organization or subject.

**EAD metadata crosswalk:** In 2002, the Research Library Group (RLG) issued a set of guidelines for EAD usage. These guidelines were designed to help facilitate uniformity and interoperability. Three useful guidelines worth noting are: “1. To facilitate interoperability of resource discovery by imposing a basic degree of uniformity on the creation of valid EAD encoded documents and to encourage the inclusion of elements most useful for retrieval in a union index and for display in an integrated, cross-institutional, setting” (RLG, 2002, p. 1).

“2. To offer researchers the full benefits of XML in retrieval and display by developing a set of core data elements to improve resource discovery…by identifying core elements and by specifying “best practice” for those elements… 3. To contribute to the evolution of the EAD standard by articulating a set of best practice guidelines suitable for inter-institutional and international use” (RLG, 2002, p. 1).
EAD Finding aid review: In order to effectively review this schema, I briefly reviewed the front-page content for EAD encoded finding aids from the following four institutions:

a. The Online Archive of California, EAD encoded finding aid entitled, “Inventory of the Department of Aging,” (see Figure #3, pg. 4).

b. Yale University Library, EAD encoded finding aid entitled, “Guide to the 250th Anniversary” (Figure #5, upper right),

c. University of Washington Library, Special Collection, EAD encoded finding aid, entitled “Guide to the AFL-CIO Region 9 Records,” (Figure #6, middle right),

d. Utah Academic Library Consortium, Mountain West Digital Library, EAD encoded, “Inventory of the Utah Women’s Democratic Club records,” (Figure #7, lower right).

In my brief examination of the format and layout of these four EAD encoded finding aids, it was apparent to me that while all of these finding aids are encoded in EAD, the front-page web layout used by each institution may be creating the appearance of a lack of uniformity.

The EAD metadata is designed to support the indexing and retrieval of the collection’s finding aid. In researching this standard, I became aware of some variations or differences in the viewable collection information. If you look at the collection titles and date spans, you will see that the collection information appears to have some variation in layout and title headings from institution to institution. If institutions presented their front-access (website) data in a standardized format, these variations would be less obvious.
Functions: In reviewing the EAD standard, the EAD metadata schema uses meta-tag elements and is used to describe a collection as a whole, to encode a detailed multi-level inventory of the collection, and to enable the collection to be indexed.

EAD was developed to enable the online retrieval of encoded finding aid data, while providing a standard structure for the data (Kiesling, 2006, p. 210). EAD was designed to be flexible in its treatment of multilevel descriptions, with a minimum of required elements (Kiesling, 2006, p. 210).

The overriding goal was to “make EAD applicable for archival arrangement and description practices” (Kiesling, 2006, p. 210). As Kiesling noted, “the emergence of EAD convinced the descriptive standards community that the creation of a content standard for finding aids was crucial” (2006, p. 212).

In a typical scenario, an archivist creates an EAD finding aid, adding subject keywords. He then creates a MARC record from the EAD finding aid and this record is added to index systems such as OCLC and RLIN (NISO, 2004, p. 5).

“Best practices suggest that EAD be used in conjunction with a MARC record to describe an archival record group or manuscript collection” (University, 2009, n.p.). “Before EAD finding aids are developed, basic arrangement and description, including the preparation of a summary descriptive record in MARC format or in a local database… must be completed” (University, 2009, n.p.).

The most important meta-tag section of an EAD-encoded finding aid is the <eadheader> (Figure #8 under References). The important elements of the EAD are: title, inclusive dates, extent or cubic feet of the collection, scope and content notes, biographical or historical note, and a statement about arrangement (Tibbo & Meho, 2001, p. 66). These metadata elements are essential as they enable the collection to be identified and cataloged online.

Users: Users of EAD include anyone whom is researching materials, topic, or a subject contained in the collection. This would include researchers, academia, students, authors, etc. In fact, anyone with an interest in a particular subject or topic could wish to access an EAD encoded finding aid.

Challenges: In my readings on EAD, I learned of three major barriers to EAD implementation: the lack of enough staff, the lack of IT software support, and the desire of archivist’s to rewrite their finding aids before implementing EAD (Yaco, 2008, pp. 471-2). Another challenge to EAD usage is the lack of cataloging skills on the part of library staff (Yaco, 2008, p. 470).

As Yaco noted, “…EAD is a technology dependent standard. A significant gap still exists between the technological expertise needed to implement EAD and the computer skills of many archivists” (2008, p. 473). The main obstacles to EAD implementation appear to be the multiple technical steps required. “One must establish…a workflow, establish the desired coding standard, encode finding aids, develop the XML, and upload everything to a web server” (Yaco, 2008, p. 457).

In summary, it appears from my brief review of the four EAD collections and the EAD readings that this metadata standard has been largely accepted by the archival community. This schema appears to have effective meta-tags to facilitate crosswalk applications. Some readings suggest that implementation of this metadata standard may be challenging. Archives or libraries, depending on their size, may be dealing with problems implementing EAD due to obsolete or insufficient finding aid descriptions, insufficient staffing, and a lack of staff cataloging expertise. But, overall, it appears that the EAD schema is an effective standard for libraries and archives to catalog their finding aids online.
References:


http://www.aip.org/history/nbl/findingaids.html

Combs, M., & OCLC Research (2010). Over, under, around, and through: Getting around barriers to


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


https://research.wsulibs.wsu.edu/xmlui/bitstream/handle/2376/1076/italNativeXml.pdf?sequence=1

[DCMI], Dublin Core Metadata Initiative (2013). Dublin Core Metadata Initiative, retrieved


10/11/2013 from: http://fclaweb.fcla.edu/Opening_Archives


Retrieved 10/11/2013 from:


http://www.tandfonline.com/doi/pdf/10.1300/J201v03n02_15


10/12/2013 from:


Bethesda, MD. Retrieved 10/13/2013 from:


Figures:

Figure #1

Figure#2
University of Nebraska, (2013). What is EAD?, Center for Research in the Humanities, retrieved 10/12/2013 from: http://cdrh.unl.edu/articles/guide_site/ead.php

Figure#3
Online Archive of California, “Inventory of the Department of Aging Records,” California State Archives, retrieved 10/12/2013 from: http://www.oac.cdlib.org/findaid/ark:/13030/kt529027dh/

Figure #4
Center for Digital Research in the Humanities, Example of an EAD finding aid template, retrieved 10/11/2013 from: http://tinyurl.com/nxp52s5

Figure #5

Figure #6

Figure #7
Figure #8

Sample EAD header.

Drexel Honesty blurb

At the beginning of each term, students must submit the following statement with their signature:

I certify that:

• I will submit only my own original work, created without unauthorized collaboration.
• I will not quote the words of any other person from a printed or audio source or a website without indicating what has been quoted and providing an appropriate attribution/citation.
• I will not submit work in this course that has been used to satisfy the requirements of any other course or created to meet other, non-academic requirements.

Signature  _______james gross______________

Date  _______10/13/2013______________