Encoded Archival Description (EAD)

Finding Aids
- Archival finding aids are tools that describe unpublished collections of personal papers and organizational records.
- Although finding aids take many forms, a more standardized structure has emerged over the past decade or so.
- The typical finding aid contains basic identifying elements such as name of creator, title, and date range. It also provides background information about the organization or person who created the records, a note on the content and formats included in the collection, statements concerning copyright and other possible restrictions, and sometimes access points like name and subject headings.
- Next comes an inventory that describes the components of the collection in more detail.
- As useful as archival finding aids are, their availability has been limited: Some finding aids have been published or presented on the Web as HTML documents, but most are available only in the repositories that generated them.

Finding Aids: Inventories
- Also called registers, have become the professional standard for describing archival collections.
- Consists of a general narrative description of the entire collection followed by more detailed summary descriptions of component series, if any.
- Following this comes a listing of the containers – boxes and folders – presented in a way that reflects their intellectual order and container type.

Finding Aids: Calendars
- Typically provide a chronological list of documents, frequently with an abstract noting type of document, creator, date and place of creation, the page or leaf count, and a summary of the contents.
- Researchers find calendars highly valuable, but they require an enormous amount of effort to produce.

Finding Aids: Lists
- Lists disclose an institution’s holdings of a well-known type of material, such as a list of holdings of newspapers, annual reports, etc.
- Lists may describe the contents of an entire collection, for example, a list of images in a photograph collection, or a portion of a collection, such as the images in an otherwise largely manuscript collection.

Finding Aids: Indexes
- Indexes provide a detailed analysis of a collection or group of related collections.
- Typically indexes include location information. For example, an institution might create an index of correspondents on a collection of the papers of a prominent person whose correspondence contains works by a wide variety of people that might interest researchers.
Finding Aids: Common Elements

- A title which the archivist has supplied
- A summary description of the material described in the finding aid
- Background and context of the collection, its creator, and the major figures involved
- Information about the custodial history of the collection, and any restrictions and conditions that apply to its use

Finding Aids: Administrative Control

- Finding aids give the location of collections on the repository’s shelves
- Finding aids identify the source, or provenance, of the collections. This provides an important link in the “chain of custody” should it ever be necessary to document for legal reasons the administrative history of a collection
- Finding aids outline the general contents of collections so that archivists can provide reference services from the moment the collection reaches the repository

Finding Aids: Intellectual Control

- Finding aids sketch the general nature of a repository’s holdings. Is this the archives if an art museum or a special collection focusing on urban affairs?
- Finding aids identify the general contents of individual collections. What are the titles of all the collections and what do they contain?
- Finding aids offer researchers detailed information about individual collections. What is in each box of the collection?
- Finding aids summarize information on a specific topic available in several collections. For example, what materials does the repository contain that would be useful for genealogical research?

Finding Aids: Administrative Control

- Finding aids give the location of collections on the repository’s shelves
- Finding aids identify the source, or provenance, of the collections. This provides an important link in the “chain of custody” should it ever be necessary to document for legal reasons the administrative history of a collection
- Finding aids outline the general contents of collections so that archivists can provide reference services from the moment the collection reaches the repository

Five Characteristics of a Good Finding Aid

- Intended for the researcher, not for the edification of the archivist; a finding aid must help the researcher to find materials
- Objective about the collection
- Aware of the needs of a wide variety of researchers
- Clear, concise, and consistent; avoids jargon
- Efficient; presents maximum usable information in a minimum of space

What is EAD?

- Defines the structural elements of finding aids and their relationships
- Expresses those concepts in the syntax of XML
  - What is XML?
    - Standard for electronic encoding of text. Facilitates computer manipulation of documents
    - Metalanguage: generalized grammar for mark up that is implemented in other more specific standards like EAD
    - Based on SGML

EAD and Finding Aids

- Encoded Archival Description is a non-proprietary standard for delineating the structural parts of a finding aid with defined SGML or XML markup tags that are embedded throughout the document.
- What those tags are and how they may be used in a finding aid is outlined in the EAD DTD (Document Type Definition), a set of rules and definitions against which an encoded finding aid is checked.
- The resulting electronic finding aid is both human- and machine-readable, can be displayed on the Web or formatted for printing, and may be saved as a plain text file.
- Because they share a standard structure, EAD-encoded finding aids from one or from many different repositories may be cross-searched in a database
Constructing a Finding Aid

Required tools: EAD Tag Library and Guidelines; DACS; LCSH; other subject headings or thesaurus (if used); local processing guide (if available)

Tips on finding aid construction:
1. Should be at collection, not item-level
2. Should be brief, not exhaustive
3. Should make evident the relationship of the records to the function or activities that created them
4. Arrangement of the finding aid should parallel that of the collection
5. Detail only exceptionally important material

Elements of a Finding Aid

1. Title page (name, volume, dates, provenance, access, copyright)
2. Administrative information (accession, preservation…)
3. Historical note (agency history/bio)
4. Scope and content note
5. Organization and arrangement
6. Controls on access
7. Description or list of series and sub-series
8. Container list
9. Item listing or index
10. Subject headings (if used)

Impact of MARC AMC

- Helping to standardize archival description
- Making information about holdings more widely known
- Sharing administrative and collections information between repositories
- Bringing the archival profession closer to related professions
- Integrating alternative access points
- Providing integrated access to unpublished and published materials

ENCODED ARCHIVAL DESCRIPTION (EAD)

- Soon the worldwide web created demand for publishing of repositories’ finding aids on websites.
- To accommodate this activity, EAD was created.
- The EAD Document Type Definition (DTD) is a standard for encoding archival finding aids using the Standard Generalized Markup Language (SGML).
- The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress (LC) in partnership with the Society of American Archivists.
- Development of the EAD DTD began with a project initiated by the University of California, Berkeley, Library in 1993.

- The goal of the 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.
- The project directors recognized the growing role of networks in accessing information about holdings, and they were keen to include information beyond that which was provided by traditional machine-readable cataloging (MARC) records.
### Characteristics of Encoding Standards

- Class of documents (archival inventories)
- Identifiable set of common elements
- Codification in a standard
  - MARC 21 standard
  - EAD DTD and Tag Library
- Standards maintenance process

### What is EAD?

EAD is:
- A data structure standard for finding aids – not a content standard
- A tool that allows finding aids to be indexed, searched, retrieved, and navigated locally or on the Internet
- Based on archival description and technical standards. EAD data is, therefore future proofed

### EAD Maintenance

- EAD belongs to SAA and the international archival community
- EAD Working Group formed in Sept. 1995: oversees intellectual maintenance
- Library of Congress Network Development/MARC Standards Office maintains official website and online documentation

### What is EAD?

EAD is:
- Helps ensure archival description will last throughout software and hardware migrations
- It is flexible enough to work with different descriptive practices

### What is EAD?

EAD is:
- Flexible enough to deal with all types of finding aids: single or multi-level, long or short, lists or calendars etc
- Can be used to create new finding aids as well as a tool for the conversion of old ones to standardized form
The Possibilities of EAD

- Increased access to archival materials through Web access
- Union databases for searching collections housed around the world
- “Virtual” collections of materials connected by provenance but dispersed in many repositories e.g. Walt Whitman Archive (University of Nebraska-Lincoln and the University of Virginia)

ENCODED ARCHIVAL DESCRIPTION (EAD)

- EAD was designed to accommodate the hierarchical structure that archivists impose on large collections of unpublished material.
- The inventory section of a finding aid begins by describing the major groups into which a collection has been organized.
- For example, a collection might be organized at the broadest level into series such as “Correspondence,” “Financial Records,” “Writings,” and “Clippings.”

Separating Markup and Display

- Content
- Presentation
- Output

- EAD document
- Stylesheet file
- Browser, Print

Example of Hierarchical Arrangement

- Each of these broad categories might be broken into smaller subseries such as “Incoming Letters” and “Outgoing Letters.”
- Subseries may then be divided into files (“Outgoing Letters A-M”).
- Here’s how a broad Correspondence series, along with its narrower subseries level and still narrower file level, would be displayed as part of an EAD-encoded finding aid:

  - CORRESPONDENCE
    - Incoming Letters
      - A-L
      - M-Z
    - Outgoing Letters
      - A-F
      - G-N
      - P-W

ENCODED ARCHIVAL DESCRIPTION (EAD)

- The principal investigator for the Berkeley Project developed requirements for the encoding standard, which included the following criteria:
  1) ability to present extensive and interrelated descriptive information found in archival finding aids,
  2) ability to preserve the hierarchical relationships existing between levels of description,
  3) ability to represent descriptive information that is inherited by one hierarchical level from another,
  4) ability to move within a hierarchical informational structure, and
  5) support for element-specific indexing and retrieval.

ENCODED ARCHIVAL DESCRIPTION (EAD)

- One problem emerged—manually produced finding aids needed to be reengineered to work properly in the web environment.
- The primary problems were that finding aids were entirely too individualistic in their creation. Elements did not appear in order or with any certainty. Thus, encoding required reconstruction of the finding aid to make sure it fit the proper mold.
What EAD enables

- Use of finding aid content as data, not just a text file
- Standardized data exchange
- Data longevity and migration
- Structured display, navigation, searching
- Multilevel description

EAD

- The Analogy
  - EAD is to archival inventories what MARC is to catalog records
- Big Concepts
  - International community standard compliant with ISAD(G)
  - Focus on content and structure rather than presentation
  - Reuse of data

XML Concepts

- Document Type Definitions/Schemas
  - Defines document structure
- Elements
  - Informational units
- Attributes
  - Modify elements
- Entities
  - External files
- Style sheets
  - Prescribe presentation

Elements (nouns)

- Have start tags and end tags
  - \texttt{<title>Moby Dick</title>}
- Have formal names and tag names
  - Formal name = paragraph
  - Tag name (generic identifier) = \texttt{<p>}

Elements (nouns)

- May contain text
  - PCDATA (parsable character data)
    - \texttt{<title>Moby Dick</title>}
- May be empty
  - \texttt{<lb>}</lb> (line break)
  - XML syntax = \texttt{<lb/>} (empty element syntax)

Elements (nouns)

- May contain other elements
  - Wrappers
  - Nesting
  - Example:
    - \texttt{<date>}
      - \texttt{<month>September</month>}
      - \texttt{<day>12</day>}
      - \texttt{<year>1958</year>}
    - \texttt{</date>}
Attributes (adjectives)

- Modify the meaning of elements
  - `<car>Honda</car>`
- Attributes of cars
  - Color
  - Year
  - Model
- `<car color="green" model="Civic" year="1996">Honda</car>`

Entities

- A set of characters references as a unit
  - Special characters
    - Language keyboard
    - Character map: XML software
    - Character entity: $141; &amp
  - Non-text files (images, sound files)
  - External data files

Style Sheets

- Separate file
- Controls presentation of data
  - Text format: font, size, color
  - Text layout: tabs, indents, line spacing, line breaks, tables
- Can supply default text and images

EAD 2002 Tag Library

- TL Conventions
- Attributes
- EAD elements
  - Brief description of element
  - May contain/may occur within
  - List of attributes and attribute values
  - Encoded examples
- Crosswalks

EAD v1.0 Application Guidelines

- Introduction and overview
- Ch. 1. EAD in Context
- Ch. 2. Administrative Considerations
- Ch. 3. Creating Finding Aids in EAD
- Ch. 4. Authoring EAD Documents
- Ch. 5. Publishing EAD Documents
- Ch. 6. SGML and XML Concepts
- Ch. 7. EAD Linking Elements

EAD v1.0 Application Guidelines

- Appendices
  - A. Minimum Recommended Elements
  - B. Crosswalks (updated in 2002 TL)
  - C. FAQs
  - D. Implementation Checklist
  - E. Examples (updated in 2002 TL)
  - F. and G. Glossary and Bibliography
EAD Design - Multilevel Description

- General description of entire body of materials (biographical sketch, scope and contents, administrative information)
- Description of individual series or other groupings (series or subseries descriptions)
- Description of files or items (container or folder list)

EAD Elements

- EAD consists of about 145 elements that describe the functions of various portions of a finding aid—elements such as <title>, <scopecontent>, <bioghist>, etc. encase the portions of information that serve that function
- It includes all 26 ISAD(G) elements
- The fact that the tag name exists within angled brackets allows the computer, and us, to distinguish between markup tags and the content of the document

EAD Elements

- Why does EAD have 145 elements vs. 26 in ISAD(G)?
  - EAD finding aid includes descriptive information about the computer file itself in the EAD Header
  - ISAD(G) is a set of core and common elements. Some elements not defined in ISAD(G) that may be necessary to describe some archival materials (e.g. bibseries)
  - ISAD(G) defines elements broadly; computers need more specific information to manipulate the content of a finding aid
  - GOOD NEWS: In most applications, only 30-40 EAD elements are used.

EAD Attributes

- Declaring the function of a piece of text with an element may not provide enough information about it
  - For example, archivists house materials in many types of containers such as boxes, folders, map drawers, etc.
  - The developers of EAD could have included a tag for each, but to avoid creating an unmanageable number of tags, they used a device called an attribute
  - Can modify generic tags like <container> by inserting an attribute and its value into the tag
  - <container type="box"><container type="folder">, etc.

EAD Elements

- <archdesc level="collection" relatedencoding="MARC">
  <accessrestrict encodinganalog="506">
    <head>Access</head>
    <p>The collection is open for research use, with the exception of the correspondence files in Series 1, Box 7, which are restricted until 2030.</p>
  </accessrestrict>
</archdesc>

Stylesheets

- Robust markup languages define the function of a text: font, color, italics, etc.
- Can create consistent presentation of presentation of all finding aids
EAD Structure

EAD splits the finding aid into segments, basically a metadata record that describes the content of the finding aid, followed by the content of the finding aid, which describes the collection. The structure then has these components:

<ead>
<eadheader>
<frontmatter>
<archdesc>
</ead>

The header creates a Dublin-core-like metadata description of the salient features of the collection (title, provenance, dates, etc.) and frontmatter describes the title-page like elements of the finding aid.

<eadheader> attributes

- audience = "external" or "internal"
- encodinganalog = "MARC"
- langencoding = "iso639-2b"
- countryencoding = "iso3166-1"
- dateencoding = "iso8601"
- repositoryencoding = "iso15511"
- scriptencoding = "iso15924"

<eadheader> EAD ID

- <eadid> unique identifier for the finding aid

  Attributes
  - countrycode
  - mainagencycode
  - url
  - urn

<eadid> examples

- <eadid>urn:taro.utexas.hrc.00001</eadid>
- <eadid countrycode="US" mainagencycode="NNC">NNCM776</eadid>

EAD Header: File Description

<filedesc>
<titlestmt>
<titleproper> <subtitle> <author>
<publicationstmt>
<publisher> <date>
</publicationstmt>
</filedesc>

Title of finding aid, not collection
Archival processor for the collection
EAD Header: Profile Description

<profiledesc>
<creation> of the electronic finding aid
<date>
PCData
<language> language of the finding aid>
<language>
<descrules> descriptive rules used (e.g. DACS)
</profiledesc>

EAD Header: Revision Description

<revisiondesc> version control
<change>
<date> <item>
</revisiondesc>

Front Matter

<frontmatter>
<titlepage>
<author>
<date>
<edition>
<publisher>
<sponsor>
<subtitle>
<titleproper>
<div>
</frontmatter>

Components of an EAD-encoded Finding Aid

- <ead>
  - Encloses all other elements and indicates to a computer that what
    follows is a document encoded using EAD
- <eadheader> (required)
  - contains metadata about the title, author, and creation date of the
  finding aid, as well as information about the language in which the
  finding aid is written and details about its encoding
- <frontmatter> (optional)
  - includes subelements that allow the creation of title pages and other
    publication-type prefatory material such as acknowledgments and
    introductions
- <archdesc>
  - contains unfolding hierarchical levels of description by first allowing a
    descriptive overview of the whole, followed by more detailed description
    of the parts, or Components <c>

EAD Structure

Archdesc then contains the whole finding aid, in essence
  describing the archival collection, using the following elements where applicable:
  
  - <did> brief descriptive identification of unit
    - <unittitle> title
    - <unitdate> inclusive and bulk dates
    - <physdesc> physical description; extent
  - <bioghist> biography/history
  - <scopecontent> scope and content
  - <controlaccess> controlled access
  - <accessrestrict> restrictions on access
  - <acqinfo> immediate source of acquisition
  - <dsc> description of subordinate components

<Archdesc>

- Wrapper element
- Must have a LEVEL attribute that identifies the highest
  level of description represented in the finding aid, e.g.,
  collection, record group, series, or subseries
- Setting the AUDIENCE attribute to "external" makes all
  the text in <archdesc> viewable to all readers
- If there is a MARC record for the collection, set the
  RELATEDENCODING attribute to "marc"
- The LANGMATERIAL attribute specifies the language of
  the collection materials
Example:
<archdesc level="collection" langmaterial="ger eng">
The high-level <did>

- The elements available within the Descriptive Identification <did> element provide basic building blocks for any level of description.
- <did> is available at all levels of description, although specific component elements may not be required at all levels (e.g., origination or repository data).
- <did> elements include:
  - Repository <repository>
  - Origination <origination>
  - Title of the unit <unittitle>
  - Date of the unit <unitdate>
  - Physical Description <physdesc>
  - Abstract <abstract>
  - ID of the Unit <unitid>
  - Physical Location <physloc>
  - Language <langmaterial>

Example of <did>

<did>
<head>Descriptive Summary</head>
<unittitle label="Title">Wolfgang Iser Papers, <unitdate type="inclusive">1952-1995</unitdate></unittitle>
<unitid label="Collection number">MS-C06</unitid>
<origination label="Creator">Iser, Wolfgang</origination>
<physdesc label="Extent">2.4 linear feet (6 boxes)</physdesc>
<repository label="Repository">University of California, Irvine. Library. Special Collections and Archives.</repository>
</did>

Repository

<repository> name of holding institution
  <corpname>
  <subarea>
  <address>
  PCData
  </repository>

<repository> example

<repository>
  <corpname>Columbia University Libraries
  <subarea>Rare Book and Manuscript Library</subarea>
</repository>

Origination

<origination> creator or collector
  <corpname>
  <famname>
  <persname>
  PCData
  </origination>

<origination> example

<origination>Provenance, William Fonds, 1897-1938</origination>
<persname encodinganalog = "100">Provenance, William Fonds, 1897-1938</persname>
Unit Title

<unittitle> title of the material being described
  <corpname><famname><persname>
  <genreform>
  <subject>
  <title>
  <unitdate>
  PCData
</unittitle>

Unit Date

<unitdate> date(s) of the unit being described
Attributes--
type (inclusive or bulk)
  normal
  standard form of the date to facilitate searching (YYYYMMDD format)

<unittitle> and <unitdate> examples

<unittitle>William Fonds Provenance Papers, </unittitle>
<unitdate type="inclusive"
  normal="1917/1937">1917-1937</unitdate>

Physical Description

<physdesc>
  <dimensions>
  <extent>
  <genreform>
  PCData
</physdesc>

<physdesc> examples

<physdesc>100 boxes (50 linear feet)</physdesc>
<physdesc><extent>10</extent><genreform>photographs</genreform></physdesc>

Identifier of the Unit

<unitid> a unique identifier for the collection, series, file, item, etc.
Attributes--
countrycode: ISO code for the country of the repository that is the custodian of the materials
repositorycode: ISO code for the repository
Abstracts and Notes

<abstract> a short statement about the contents and creator of the materials to assist users in selecting relevant collections

<note> explanatory text, requires a <p> subelement, is also available outside <did>

<note> and <abstract> examples

- <note><p>To request materials, please put the collection name, and the box and folder numbers on your call slip</p></note>

- <abstract>The papers of William Fonds Provenance, author of the landmark <title render="italic">Quarks: The Tao of Archives</title> document his college years …</abstract>

Language of the Material

<langmaterial>
  <language>PCData</language>
</langmaterial>

Container

<container>
  <container type="Box">1</container>
  <container type="Folder">5</container>
  <container type="Reel">7-10</container>
</container>

<langmaterial> example

Correspondence in
  <language>French</language>, <language>German</language> and <language>English</language>
</langmaterial>
Physical Location

<physloc> physical location within the repository

<physloc type="repository">5C:SW:6.5</physloc>

<physloc>Collection is in offsite storage. Requires 24-hour notice for paging.</physloc>

Material Specific Details

<materialspec>
<num>
<title>PCData</title>
</num>
</materialspec>

More often used at item level unless entire collection is VHS, etc.

<materialspec> example

<materialspec>Scale<num type="Scale">1:50,000</num></materialspec>

<materialspec>Duration 1:47</materialspec>

Descriptive Identification of the Unit

<archdesc>
<did>
<repository>
<origination>
<unittitle>
<unitdate>
<physdesc>
<unitid>
<abstract>
<langmaterial>
<materialspec>
<physloc>
<container>
<note>
<dao> and <daogrp> digital archival object <head>

Biography or History

• Contextual information about the creation or formation of a body of archival materials in the form of either a biographical sketch or an agency history
• Information can be in the form of either narrative text or a chronology in which dates and events are paired in columnar form

Biography or History

<archdesc>
<bioghist>
<head>
<p>
<chronlist> chronology
<chronitem>
<date>
<event> or <eventgrp>
<bioghist> (recursion)
<dao> e.g. picture of creator
</bioghist>
Example of `<bioghist>`

```xml
<bioghist>
  <head>Biographical Chronology</head>
  <chronlist>
    <chronitem>
      <date>1926</date>
      <event>Born on July 22 in Marienberg (Saxony), Germany.</event>
    </chronitem>
    <chronitem>
      <date>1946-1946</date>
      <event>Student at the University of Tubingen.</event>
    </chronitem>
    ...
  </chronlist>
</bioghist>
```

Scope and Content

- Summarizes the range and topical coverage of the described materials, often mentioning their form and organization, and naming significant individuals, organizations, events, places, and subjects represented in the materials.
- Can contain:
  - `<arrangement>`

```xml
<archdesc>
  <scopecontent>
    <head>Collection Scope and Content Summary</head>
    <p>The Wolfgang Iser Papers contain material related to his publications, reading and lecture notes for courses taught throughout his career, and papers presented at conferences. The papers reflect Iser’s main interests as a professor of English literature and as a leading literary theorist. The bulk of the material dates from after his appointment in 1967 as a professor at the University of Konstanz, in Germany; some material documents his teaching at the University of California, Irvine from 1978-1995.</p>
  </scopecontent>
</archdesc>
```

Arrangement

```xml
<archdesc>
  <arrangement>
    <p>The collection is arranged in three series ...</p>
  </arrangement>
</archdesc>
```
Controlled Access
<controlaccess>
- To provide maximum access to the materials through authority-controlled searching, the Controlled Access Headings <controlaccess> wrapper element allows the use of terms analogous to those used in MARC 1XX, 6XX, and 7XX fields.
- Sub-elements within <controlaccess> include <persname>, <corpname>, <genreform>, <occupation>, <famname>, <geogname>, <subject>, and <title>, which can be used along with [source] attributes specifying the controlled vocabulary being used, i.e., "lcnaf", "aat" or "lcsh."
</controlaccess>

<controlaccess> examples
<controlaccess>
<persname source="lcnaf" encodinganalog="100">Clinton, Bill, 1946-</persname>
<title style="display: inline">Controlled Access</title>
</controlaccess>

Administrative Information: May use <descgrp> as wrapper element
<archdesc>
<accessrestrict> conditions governing access
<userestrict> conditions governing use
<custodhist> custodial history
<altformavail> alternate forms of materials
<originalsloc> location of originals
<phystech> physical and technical specs
<prefercite> preferred citation
</archdesc>

<accessrestrict> example
<accessrestrict>
<head>Availability restrictions:</head>
<p>Until 2015, access to the collection requires written permission. Contact reference staff for more information</p>
</accessrestrict>
<custodhist> example

<p>Following his death, Provenance’s papers were bequeathed to his brother, Harry, who attempted several times, unsuccessfully, to sell them to a private collector. Upon Harry’s death, the papers became the property of Provenance’s longtime companion Ima Gusdorf, who donated them to the Freen University Center for the Study of Giants of Archivy.</p>
</custodhist>

<altformavail> and <originalsloc> examples

<p>Diary available on microfilm.</p>
</altformavail>

<p>Photocopies only available. Originals still in custody of the donor.</p>
</originalsloc>

<p>Many of the prints show fading and silvering.</p>
</phystech>

<p>Please cite as: Rick Block Papers, Manuscript Division, Oxford University.</p>
</prefercite>

More Administrative Information

<acqinfo> acquisitions information
<accruals>
<appraisal>
<processinfo> processing information
<acqinfo> example

<acqinfo>
<head>Acquisition</head>
<p>Acquired as a gift from Ima Gusdorf, November 15, 1987.</p>
<p>Accession Number: 1987:56</p>
</acqinfo>

Example of administrative information

<descgrp>
<head>Administrative Information</head>
<accessrestrict>
<head>Access</head>
<p>Collection is open for research.</p>
</accessrestrict>
<p>Property rights reside with the University of California. Literary rights are retained by the creators of the records and their heirs. For permissions to reproduce or to publish, please contact the Head of Special Collections and Archives.</p>
</descgrp>

Example of administrative information (cont)

<prefercite>
<p>Wolfgang Iser Papers. MS-C06. Special Collections and Archives, The UC Irvine Libraries, Irvine, California.</p>
</prefercite>

<acqinfo>
<head>Acquisition Information</head>
</acqinfo>

<accruals> examples

<accruals>
<p>No additions to the Provenance Papers are anticipated.</p>
</accruals>

<accruals>
<p>Additional accruals are deposited at the end of every fiscal year.</p>
</accruals>

<appraisal> example

<appraisal>
<p>The appraisal of this collection was carried out in consultation with Robert Heinmiller, a research associate at Woods Hole Oceanographic Institution during MODE.</p>
</appraisal>

<processinfo> example

<processinfo>
<p>Collection was organized by Ima Archivist, July 8, 2005</p>
</processinfo>
**Adjunct Descriptive Data**

- `<bibliography>` append or imbed. Full range of citation elements available
- `<fileplan>` e.g. filing order created by agency
- `<index>` to collection
- `<list>` formatted documents
- `<otherfindaid>` other finding aids for the materials
- `<relatedmaterial>` material NOT related by provenance
- `<separatedmaterial>` material related by provenance

**Description of the Archival Unit**

- `<archdesc>`
- `<did>` descriptive identification of unit
- `<bioghist>` biography/organizational history
- `<scopecontent>` scope and contents
- `<arrangement>` physical/logical ordering
- `<controlaccess>` controlled access points
- `<odd>` other descriptive data
- `<descgrp>` description group
- `<dsc>` description of subordinate components

**<dsc> Subordinate Components**

- `<dsc>` description of subordinate components
- `<c01 level="series"> [description of component]`  
  - `<unititle>`
  - `<admininfo>`
  - `<bioghist>`
  - `<scopecontent>`
  - `<controlaccess>`
  - `<did>`
    - `<c02 level="subseries">`  
      - `<did>`...
    - `...`
    - `...`

**Which elements are required?**

- `<ead>`
  - `<eadheader>`
  - `<eadid>`
  - `<filedesc>`
  - `<titlestmt>`
    - `<titleproper>`
    - `<archdesc>` and LEVEL attribute
  - `<did>`

**EAD and ISAD(G)**

ISAD(G) states that to be a conformant archival description a finding aid must:
- Be hierarchical: That is it must be made up of a number of levels, and must follow the four rules of multi-level description
- Have certain data elements

EAD is specifically designed to allow the representation of ISAD(G) finding aids

**EAD and Hierarchy**

ISAD(G) levels:
- Fonds
- Sub-fonds
- Series
- Sub-series
- File
- Item

EAD levels:
- `<ARCHDESC>`
- `<DSC><C01>`
- `<C02>`
- `<C03>`
- `<C04>`
- `<C05>`
ISAD(G) to EAD

**ISAD(G)** (v.2)
- 3.1.1 Reference code(s)
- 3.1.2 Title
- 3.1.3 Dates of creation
- 3.1.4 Level of description
- 3.1.5 Extent of the unit
- 3.2.1 Name of creator
- 3.2.2 Administrative/Biographical history
- 3.2.3 Custodial history
- 3.2.4 Immediate source of acquisition
- 3.3.1 Scope and content
- 3.3.2 Appraisal, destruction and scheduling

**EAD 2002**
- `<unitid>` countryCode and repositorycode attributes
- `<archdesc>`
- `<unittitle>`
- `<unitdate>`
- `<archdesc>` and `<c>` level attribute
- `<physdesc>`, `<extent>`
- `<origin>`
- `<bioghist>`
- `<custodhist>`
- `<acqinfo>`
- `<scopecontent>`
- `<appraisal>`

Hierarchical elements
- `<archdesc>`
  - `<dsc>`
    - `<c>`
      - `<c01>`...
      - `<c12>`
    - or
    - First component to twelfth component

Presentation elements
- **Headings**
  - `<head>`
  - `<p>; <lb>`
- **Layout**
  - `<emph>;`<blockquote>
- **Italics and quotes**
  - `<list><item>`
  - `<chronlist><chronitem>`
- **Lists**
  - `<ref>; <ptr>; <dao>`
- **References, pointers and links**
  - `<bibliography>`
  - `<odd>`

**Normalization of Dates**
- Normalize all dates according to the ISO 8601 standard where possible
- Remember that normalized dates are not displayed. They are used to support information retrieval queries based on dates. Only the element content is displayed.

**Examples**
- Date spans:
  - `<unitdate normal="1956-01/1956-07">Jan 1956 - July 1956</unitdate>`
  - [use ISO 8601 date intervals]
  - `<unitdate normal="1900/1950">1900-1950</unitdate>`
- Broken date spans (e.g., "1924, 1956-1975"):  
  - [encode dates in separate `<unitdate>` tags]
Examples

- Open date spans:
  - `<unitdate normal="1911/9999">1911-[ongoing]</unitdate>`
  - [use an interval and set the end date to 9999]
- Approximate dates (e.g., “ca. 1950”):
  - `<unitdate normal="1945/1955">circa 1950</unitdate>`
  - [normalize as an interval to express an appropriate date range]
  - `<unitdate normal="1980/1989">1980s</unitdate>`
  - [use an interval to indicate every year of the decade]
  - `<unitdate normal="1801/1900">19th century</unitdate>`

Examples

- Undated material:
  - `<unitdate normal="1920/1957">undated</unitdate>`
  - [normalize as an interval (as with approximate dates), perhaps using the collection dates, or life of creator, etc.]
- Undated material:
  - `<unitdate normal="1935/1965">undated: circa mid 20th century</unitdate>`
  - [if a document is undated this can be stated but provide an estimate if possible; normalize as an interval, perhaps using the collection dates, or life of creator, etc.]
  - `<unitdate>undated</unitdate>`
  - [if it is really not possible to estimate then use “undated” alone with no normalization]

Special Characters

- EAD documents encoded in XML use the UTF-8 Unicode character encoding format. When using XML, characters used as markup delimiters must be replaced by character entities, as in the following table:

<table>
<thead>
<tr>
<th>Character</th>
<th>Name</th>
<th>Entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>&amp;</td>
<td>ampersand</td>
<td>&amp;amp</td>
</tr>
<tr>
<td>&lt;</td>
<td>left angle bracket</td>
<td>&lt;</td>
</tr>
<tr>
<td>&gt;</td>
<td>right angle bracket</td>
<td>&gt;</td>
</tr>
</tbody>
</table>

Advantages of EAD

- Created by archivists for archivists
- Compatible with ISAD(G)
- Stable, non-proprietary
- XML is THE international standard for data exchange
- Electronic delivery
- Designed to represent complex hierarchies

Advantages of EAD

- Multiple uses (repurposing)
- Powerful XML tools for searching, retrieving, displaying and navigating
- Long-term accessibility
- Cross-searching: advantages of being part of larger community
- EAD supports ISAD(G) element set completely, particularly multi-level description
- EAD was created by archivists, for archivists
Caveats

- Complex suite of technologies
- Requires evaluation of hardware and software options
- Ongoing maintenance
- Technical expertise
- Training
- Commitment of resources
- Changes to workflow methodologies?

Implementing EAD: Creating Encoded Finding Aids

- Adopt an encoding protocol
  - Ensure consistent markup
  - Enable interoperability
  - Facilitate display
  - Consortial protocols
  - RLG Best Practices Guidelines for EAD

Implementing EAD: Creating Encoded Finding Aids

- Select, install and configure authoring software
  - EAD Cookbook
  - Templates
- Encode finding aids
  - New documents
  - Covert existing finding aids
  - Rokey text
  - Cut-and-paste into an editor
  - Outsource markup
  - MARC 21

Conversion of finding aids

- priorities
  - heavy/low use
  - added value from online retrieval
  - 'unite' split holdings
  - link to digital resources
  - funding opportunities
- how much editing is required?
- current formats?
- outsource?

Conversion: Encoding Procedure

- Retyping – in-house or data conversion company
- Put into EAD template
- Scanning and OCR
- Scripts to automate (part of) the process
- Conversion programmes

Conversion of finding aids

- “Before we try to convert our finding aids into EAD-encoded documents we ought to make certain that those finding aids are as well thought out as possible in terms of both their structure and their content.”--Dan Meissner
- Common problems plaguing existing finding aids
  - Information elements not clearly identified. This makes it difficult to determine where to place the EAD tags
  - Information elements are not optimally arranged. Even if all elements are present, they may be in an order that will make research difficult
  - There can be inconsistent levels of description. Various series may be described to different levels. This inconsistency is a problem when converting to EAD
  - User instructions may be lacking. The finding aid may not help the researcher to find his or her way through the finding aid or collection
New Finding Aids
- standardization
- authority records (LCNAF, LCSH)
- appropriate software
- quality control
- staff expertise
- EAD is not about display
- Style sheets are used for formatting

Implementing EAD: Enable Resource Discovery
- Full-text databases
- Web search engines (Google, Yahoo)
- Links from MARC records
  - Field 856 Electronic location and access
  - Field 555 Finding aid availability note
- Subject-oriented web pages

Implementing EAD: Deliver Finding Aids on the Web
- Adapt or adopt existing stylesheets or create new ones
- Transform finding aids into HTML
- Mount finding aids on web server

What are Stylesheets?
- XSLT stylesheets are computer programs that tell XSL processor software how to transform an XML file
  - Control the display of the EAD text
    - Text format: font, size, color
    - Text layout: tabs, indents, line spacing, line breaks, tables, page breaks
  - May insert text and images into the output file
    - Institutional logo, address, phone numbers
    - Images relating to the collection
    - Images of collection materials
- For finding aids, we transform an EAD document into an HTML file for viewing in a browser on the web
- Stylesheets may be created or modified with specialized software or any text editor
Creating Stylesheets: What the Archivist Needs To Do

• Review current finding aids
  – Data content
  – Order of information
  – Layout of text
• Determine new presentation
  – Table of contents
  – Text display
  – Order of data

Creating Stylesheets: What the Archivist Needs To Do

• Create encoding that supports the desired output
  – Heads and labels
  – Hyperlinks
• Apply encoding consistently
  – Encoding protocol
  – Templates

Where to get files and help

• http://www.loc.gov/ead/
  – Official EAD website with online Tag Library and Application Guidelines
• http://www.archivists.org/saagroups/ead/
  – SAA EAD Roundtable Help Pages—Ead Cookbook templates in XMetaL, NoteTab and <oXygen>, and stylesheets
• Library of Congress EAD listserv
  – ead@loc.gov
  – listserv@loc.gov (subscribe ead your name)

Encoded Archival Context – Corporate Bodies, Persons, and Families (EAC-CPF)

• EAC-CPF is a communication structure for archival contextual information for individuals, corporate bodies and families.
• It supports the exchange of ISAAR (CPF) compliant authority records. ISAAR (CPF) "determines the types of information that could be included in an archival authority record and provides guidance on how such records may be deployed in an archival descriptive system."
• ISAAR (CPF) also notes that "successful automated exchange of archival authority information over computer networks is dependent upon the adoption of a suitable communication format by the repositories involved in the exchange."

EAC-CPF

• EAC is one such communication format which supports the exchange of ISAAR(CPF) compliant archival authority data over the World Wide Web."
• EAC-CPF is designed to be, among other things, an implementation of the International Standard Archival Authority Record for Corporate Bodies, Persons, and Families -ISAAR(CPF)
Reimagining Description and Access (1)

• Rigorous analysis of the logic and structure of archival description
• Recognition of the functional inadequacy of single apparatus
• Increasing differentiation and formal definition
  – Components of archival description
  – Relations between components

Reimagining Description and Access (2)

• Reimagining not entirely new
• Peter Scott, "The Record Group Concept: A Case for Abandonment" *American Archivist* 29:493-504 (October 1966)
• Advanced technologies: means to realizing description and access

Components of Archival Description

• Description of records (as such)
• Context: creators
• Context: functions and activities documented in records
• Components interrelated with one another
• Dedicated descriptive semantics and structure for each component

What Needs to Be Done

• Complete work on EAC-CPF and EAC-F
• Revise EAD
  – Accommodate moving the present into the future
  – EAD needs to accommodate EAC-CPF and EAC-F namespaces
  – Simplify EAD
    • Considerably less "mixed content"
    • Move "label" and "head" to out-of-line, i.e., make `<archdesc>` purely about description
• Make EA… standards relational database "friendly"

Current Status of Work

• EAC-CPF released Aug. 21, 2009
• Preliminary testing
  • 109 MARC records > XML Slim > EAC-CPF
  • Three Australian records
    – Bright Spars > EAC-CPF
    – People Australia (ANL)
• EAC-CPF Tag library in draft (multilingual)
• ISDF/EAC-F to follow
• EAD: revise here, revise now!
EAC-CPF

- Authority control for corporate bodies, persons, and families but more …
- Controlled vocabulary description of named entity (place, occupation … and extensible)
- Prose biography or history of entity
- Chronological list (date, place, event)

EAC-CPF

- Designed to be relational database “friendly”
- Designed to be used in an international, multilingual, and shared environment
- Designed to enable ingesting and integrating
  - Authority control records
  - Biographies and histories
  - From two or more sources
  - Based on one or more sets of descriptive rules
  - To provide union access

International Standard Archival Authority Record for Corporate Bodies, Persons and Families: ISAAR(CPF)

- ISAAR CPF provides guidance on creating archival authority records for corporate bodies, persons and families associated with the creation and maintenance of archives.
- Archival authority records may be used to describe these entities; to control the creation and use of access points within an archival description; to document relationships between the entities and the records created by them; and to link to other resources about, or by, them.

Records

- Archival records: the byproduct of people, either as individuals, families or organized groups, living their lives, or doing their jobs
- Records, broadly speaking: the artifacts of human activity: books, journals, art, natural objects collected, and so on
- Creator description important for archivists, but also everyone that describes records, in the broad sense: librarians, curators, and others

Responsibility

- Archivist have a special responsibility for developing a standard for creator description
- Essential to description and control of archival records-Context
- Archives have the primary resources on which biographies and histories are based
Standardization

- Economic benefits
  - Creator description expensive: research, analysis, creation, and maintenance
  - Records with common provenance dispersed
  - Records with related provenance or subjects dispersed
  - Complementary records and creators
  - Standards will support collaboration and sharing

Standardization

- Enhanced access and services
  - Controlled access: documentation of multiple names for the same entity
  - Biographical and historical information will enhance access and serve as a resource in itself
  - International, internet accessible biographical and historical resource?

Context

- Creator description implicitly a major component of archival description, and indeed description of resources in general
- The opportunities presented by computer technology for facilitating description and control of records inspiring new and more rigorous analysis of the logic and structure of description

Context

- Increasing differentiation and formal definition of the
  - components of archival description
  - relations between the components
- Historically description of creator embedded in the description of records
- Increasing recognition that creator description more economically done by separating creator description from record description

Context

- Increases flexibility in record description
- Recognition that creator description can play roles other than providing context for record origins
  - Important and first step in discovery of records and other creators; authority control and access
  - Important reference resource in itself; biographical and historical information

Archival Description

- Three principal entities
  - Creators
  - Functions and activities
  - Records: byproducts and products
- Interrelated
  - Place
  - Date
Relations-Creator Centric View

- Creator to creator
- Creator to functions and activities
- Creator to records (broadly defined)
- Each qualified by place and time
- Subjects, verbs, objects, date, and place
- Records are evidence of functions and activities, though can be objects

Overview of EAC-CPF Structure and Semantics

- Each EAC-CPF document contains two mandatory elements, the 
  - <control> element
  - either the element <cpfDescription> or <multipleIdentities>
- The <control> element contains data used in the control of the entity description, and to provide context for that description.
- <cpfDescription> contains information on the name structures, descriptive elements, and relationships.
- <multipleIdentities> is used when there is more than one <cpfDescription>.
- These two wrapper elements contain specific elements to support the functional intentions of the parent or containing element.

Authority Control

- Identifying creator entities
- Recording name or names used by and for them
- Rule-based heading or entry formation and control

Overview of EAC-CPF Structure and Semantics

- SINGLE IDENTITY: one person (or corporate body or family) with a single identity represented in one eac-cpf instance. (Most common)
- MULTIPLE IDENTITY-MANY IN ONE: two or more identities (including official identities) with each represented by distinct descriptions within one eac-cpf instance.

Objectives

- ICA: ISAAR(CPF): implementation and revision
- Complement EAD: <bioghist>
- Capable of importing existing authority, biographical, and historical data: MARC, print, and so on
- Original input
- Communication

<control>

- <recordId> - EAC-CPF identifier. Contains one or more unique identifiers for the EAC-CPF instance. Required.
- <maintenanceAgency> - Maintenance agency. Name and coded information about the institution or service responsible for the creation, maintenance, and/or dissemination of the EAC-CPF instance. Required.
- <maintenanceStatus> - Maintenance status. Contains the current drafting status of the EAC-CPF instance. Values include: new, revised, deleted, cancelled, deletedSplit, or deletedReplaced. Required.
- <maintenanceHistory> - Maintenance History. Contains information about the date, type, and events within the lifecycle of an EAC-CPF instance. Contains one or more <maintenanceEvent> elements that document creating, importing, updating, and deletion of the description. Each maintenance event contains an agent, the type of agent (human or machine), the type of event, a description of the event, and the date of the event. Required.
---

**<control>**

- `<sources>` - Sources. Contains information about the sources consulted in creating the description of the entity or entities in the EAC-CPF instance. Contains one or more `<source>` element. Required.
- `<ruleDeclaration>` - Rule declaration. Contains information on the rules used to construct the EAC-CPF instance, in particular the names formed in `<identity>`. Optional.
- `<authorityDeclaration>` - Authority declaration. Contains information on the controlled vocabularies and thesauri used in the EAC-CPF instance. Optional.
- `<otherRecordId>` - An element that allows the recording of additional identifies that may be associated with the EAC-CPF instance. Optional.
- `<localControlEntry>` - An element in which to record any control entries necessary due to local practice that are not represented by the other elements in `<control>`. Optional.

**<cpfDescription>**

- The `<cpfDescription>`-Corporate body, person or family description, comprises the description of the entity. Similar to the `<control>` element, `<cpfDescription>` has several complex subelements used to describe different features of the entity.
  - `<identity>` - Identity. Complex structure containing the name or names used by the entity over the course of the entity’s existence. Contains `<nameEntry>` elements for different names, and `<nameEntryParallel>` for more than one `<nameEntry>` expressed in different languages.
  - `<description>` - Description. Contains formal description elements parallel to those in ISAAR (CPF) for the description of the (agent) entity. An additional `<descriptiveEntry>` allows for local implementation of additional descriptive information not included in the other `<description>` elements.
  - `<relations>` - Relations. Contains references to or descriptions of related corporate bodies, persons or families `<cpfRelations>`, functions `<functionRelations>`, or resources `<resourceRelations>`.

**<identity>**

- The most complex element in the EAC-CPF schema is the `<identity>` element.
- In addition to needing to accommodate one or more names used for or by the entity, `<identity>` must accommodate two or more parallel names in different languages or scripts.
- In countries where there is more than one official language, such as Canada, names of corporate bodies have more than one language.
- The `<identity>` contains a required `<entityType>` and one or more `<nameEntry>` and/or `<nameEntryParallel>` elements.
- It also includes an optional `<entityId>` and `<descriptiveNote>`.

**<description>**

- The `<description>`-description element accommodates a variety of both controlled and prose descriptions of agents.
- These elements reflect the descriptive categories outlined in ISAAR (CPF).
- Descriptive elements generally have a singular and plural form, the latter being used for those cases of multiple instances of a descriptive category. For example, `<function>` would be used for a single function term, `<functions>` will bundle more than one function descriptor.
- With the exception of `<existDates>`, all description elements are optional.
- In addition to the elements specific to the descriptors, all elements within `<description>` include an optional `<descriptiveNote>` element for a fuller explanation of the descriptor.

---

**<existsDates>** (Required) – dates of existence of the entity being described. Can include actual or approximate dates, using either `<date>`, `<dateRange>`, or `<dateSet>`.

**<placeDate>** -- includes relevant location information paired with related date information. Includes elements `<place>` and the range of possibilities with date information: `<date>`, `<dateRange>`, `<dateSet>`.

**<descriptiveEntry>** -- An element intended to extend the descriptive categories available in a local system. Includes a `<term>` element and the range of possibilities with date information: `<date>`, `<dateRange>`, `<dateSet>`.

**<legalStatus>** -- Includes the legal status of a corporate body, typically defined by authorities and granted by either a government or an authorized agency.
<description>

- <function> -- Includes relevant functions performed by the entity being described. Includes a <term> element and the range of possibilities with date information: <date>, <dateRange>, <dateSet>.
- <occupation> -- Includes relevant occupations held by the entity being described. Includes a <term> element and the range of possibilities with date information: <date>, <dateRange>, <dateSet>.
- <activity> -- Includes relevant activities conducted by the agent entity being described. Includes a <term> element and the range of possibilities with date information: <date>, <dateRange>, <dateSet>.
- <mandate> -- Includes relevant mandates for corporate bodies being described. Includes a <term> element and the range of possibilities with date information: <date>, <dateRange>, <dateSet>.
- <structureOrGenealogy> -- Includes information about the structure of a corporate body or the genealogy of a person or family. Includes elements <outline>, <list>, and <p> to assist in structuring the text.
- <generalContext> -- Includes information about the general social and cultural context of the entity being described. Includes a <term> element and the range of possibilities with date information: <date>, <dateRange>, <dateSet>.
- <biogHist> -- Includes narrative text providing biographical and/or historical information about the entity being described. Includes an <abstract> element for a brief synopsis of the full content; a <chronList> element allows for structured date, event and optional place information. Includes <list>, <outline>, <p> elements to assist in structuring the text.

All elements in <description> provide a LOCALTYPE attribute to identify the controlled vocabulary or thesaurus connected to the specific term being used.

With the exception of <existDates>, <structureOrGenealogy>, and <biogHist>, plural form wrapper elements are available to bundle multiple occurrences of these elements.

These wrapper elements also include elements <citation>, <list>, <outline>, and <p> to accommodate greater complexity in representing the description being created.

<relations>

- There are three elements for describing relations with other descriptive entities that are included in the <relations> element: <cpfRelation>, <functionRelation>, <resourceRelation>.
- Within each of these relations elements, there are: <relationEntry>, <objectXMLWrap>, <relationBinWrap>, <date>, <dateRange>, <dateSet>, <place> and <descriptiveNote> elements.
- Individual relations include optional attributes related to the type of relation that is being described.

- <cpfRelation> -- CPFRELATIONTYPE; values include identity, hierarchical, hierarchical-parent, hierarchical-child, temporal, temporal-earlier, temporal-later, family, associative
- <functionRelation> -- FUNCTIONRELATIONTYPE; values include activity, ambientFunction, functions, process, transaction, other.
- <resourceRelation> -- RESOURCERELATIONTYPE; values include creatorOf, subjectOf, other