EAD Task Force meeting minutes  
June 15, 2005

1. DLXS and PLEADE  
Amy asked about people’s responses to the two search engines Sheila Yeh has set up for testing: DLXS and PLEADE. Most Task Force (TF) members prefer DLXS and find that it gets better reactions from other people they’ve shown it to. It’s easier to comprehend that you are looking at a document, rather than isolated records, and navigation is simpler. If DLXS’s upcoming new release offers full XSLT functionality as promised, then the configuration and display problems Sheila has been dealing with should be eliminated. Similarly, if PLEADE’s next release includes the increased functionality they have promised, it should be possible to make the search results more user-friendly. TF members agreed we have time to wait and see what is actually delivered, then revisit the decision on search engines.

There was a discussion of linking finding aid records to digital images. Everyone said that once their finding aids are online, it will be important to include images too, so this should be a part of the search engine development. If PLEADE is adopted, Sheila will need to obtain documentation and/or training for implementation of Navimages. DLXS promises support for JPEG 2000 in its new release.

A brief discussion of JPEG 2000 followed. Peter Murray, who discussed the format at the previous EAD TF meeting, will do a presentation about it as part of a program at the American Library Association annual meeting in June. Amanda will attend the program and report on it at the August 31 meeting. Amy will report on a JPEG 2000 presentation at the Society of American Archivists annual meeting in mid-August.

Barbara requested that her name be added to the EAD TF Web page, and asked that the starting time of each meeting be posted along with the meeting date and agenda.

2a. Content Guidelines—top-level elements  
The group reviewed an initial list of required top-level elements, suggested additions, and talked about whether certain tags should be required or merely recommended.

Specifications for certain elements were discussed:

<origination>
Names the creator of the collection. The issue of donors who don’t want to be identified was discussed. It was agreed that the “audience” attribute for this element can be set to “internal.” The same is true for names in <custodhist> and <acqinfo>.

<scopecontent> and <abstract>
An abstract summarizes a scopecontent note, especially useful when the scopecontent note is lengthy. Instead of using an abstract, one can write a succinct scopecontent note, or one that has a summary lead paragraph. TF members stated their preference for a
well-written `<scopecontent>`, beginning with a summary, over using a separate abstract. Everyone agreed that directions for writing a `<scopecontent>` appropriate for a MARC 520 field should be a part of the EAD content guidelines.

 `<controlaccess>
The group discussed whether these should be required at the top level. No decision was made; the issue will be revisited.

 `<custodhist>
Should not be required; often this information is not known. This element may occur at multiple levels if the collection is composed of more than one donation.

 `<arrangement>
Group agreed it should be required. There was some discussion of whether it should be nested in `<scopecontent>`, or be a separate element. The group would like to allow for both practices. Amy will ask Sheila to test this to see if search engines handle both cases without special configuring.

 `<physloc>
Physical location, can be a shelf, a room, a building. Group favors setting the audience attribute to “internal” because the public doesn’t need to know, and in some cases giving locations could be a security risk.

It was agreed that the EAD content guidelines should be in two parts: 1) A master list of elements, as a starting point, and 2) Separate descriptions of each element. Agreed that the RLG Best Practices chart is confusing, especially to beginners, so we will create a simpler presentation.

### 2b. Content Guidelines—DACS-defined elements

The group proceeded to review DACS sections assigned to each TF member at the previous meeting. Each member presented specifications and comments for their elements:

 `<unitid>
A unique alphanumeric string identifying a collection, often derived from an accession code. Cara pointed out that some repositories don’t use accession numbers, so many legacy collections don’t have them. It was proposed that the content guidelines include an information page explaining how to construct one. Possibilities include using a local call number, or using a repository code plus a year/arbitrary accession number (as in KSU2005.22).

Toni suggested notifying implementers that they will need a repository code, and should start thinking about accession numbers as well. This could take the form of a “Getting
Started” or “First Steps” section in the content guidelines. Step 1, get a code—Step 2, create accession numbers—Step 3, get familiar with DACS, etc.

TF members noted that some of the elements under discussion were overlapping. For instance, DACS presents a top-level <unitid> and a component-level <unitid> in two different areas, so Amanda and Cara both wound up with this element. It was emphasized that although the TF is using DACS as a starting point, ultimately the presentation of elements will be EAD-driven, not DACS-driven.

<extent>
Group agreed that “cubic feet,” “linear feet,” and “number of boxes” should all be allowed.

<physdesc>
Parent element of <extent>, has numerous child elements serving various purposes. Group discussed presentation in the content guidelines of this and other complicated elements. Agreed that the parent element page should link out to subelement pages in some cases, but not all. For instance, many elements can have <p> or <extref> as subelements, but linking to these elements repeatedly would be excessive.

<unittitle><unitdate>
Group discussed whether to require separation of these elements, in accordance with international practice, instead of allowing for <unitdate> nested in <unittitle>, a practice often used in the U.S. Group tentatively decided to require them as separate in the content guidelines, and make recommendations for dealing with legacy finding aids that present nested date problems. It may be a good idea to have a separate section of the guidelines dealing with legacy finding aids.
The “type” attribute for <unitdate> (“inclusive” or “bulk”) will be required at the top level, optional at the component level.

<container>
The EAD DTD contains a list of allowed container types. This list can be edited. The TF will recommend a list of container types to be included.

<appraisal>
An optional element, describing the details of the archivist’s appraisal of the collection. Can be set to “internal.”

There was a discussion of highly specific elements such as <phystech> and <altformavail>. Should the recommendations lean toward more generality in encoding, saving time and keeping files smaller? It was suggested that there be an explanatory note in the content guidelines stating that the use of optional elements should be carefully considered.
Cara suggested that the TF draft one basic, restrictive template, and consider allowing OhioLINK participants to submit valid alternate templates for special materials, such as photographs.
The group discussed elements and attributes that are embedded in EAD documents as code that makes machine processing of the documents possible. Encoders don’t have to do anything with them, as their input is automated. The content guidelines will make this clear, so that EAD beginners who happen to glimpse them don’t get worried or frustrated by their complexity.

Barbara pointed out that if you send your XML file to a consortium, you don’t have any guarantee that confidential information will be kept private. There is no control over how an outside agency will manipulate it. Cara pointed out that if you use EAD as a “management tool” for administrative information, you may risk exposing private donor information and other non-public info. It was suggested that the content guidelines should include a note on privacy that makes this clear.

2c. Content Guidelines—classifying elements as required or optional
The group discussed the classification of elements as required or recommended, based on what everyone knows about typical archival practice. There was a discussion of terms used in indicating whether particular EAD elements are required, required in some cases, or optional. The following designations were decided on:
Required (R)
Required if applicable (RA)
Optional (Opt)

The review of individual elements will continue at the next meeting.

3. Presentation of HTML template and output
Amy presented a sample HTML page for EAD element specifications in the content guidelines. A NoteTab template for creating these pages will be distributed to the EAD TF discussion list. Once the element specifications are completed, they can be stored in a non-public development directory on the TF web site.