EAD @ 10
THOUGHTS ON THE FUTURE

Jeanne Kramer-Smyth
http://www.spellboundblog.com
August 31, 2008
Beyond the Markup

- Tools
- Centralized Repositories and Union Catalogs
- Data
- User Generated Content
- Visualization
More Tools

- Tools for creating EAD
- Tools for conversion between formats
- Tools for visualizing EAD
Possible Impacts of Tools

• Small archives can join the fun
• Reasons to use EAD beyond ‘getting the finding aid online’
• Faster and easier to gain high level understanding of records available
• Better tools for administrators
• And eventually... improved paths for the general public to follow to primary sources
More Centralized Repositories

Examples:
• Online Archive of California (OAC)
• Northwest Digital Archives (NWDA)
• ArchiveGrid

Related efforts:
• Descriptive Metadata Guidelines
• Further work on centralized thesaurii
Metadata Guidelines

CDL Guidelines for Digital Objects: Descriptive Metadata Guidelines

• The form of the heading should be taken from a standard or local thesaurus
• If possible, indicate the code for a standard naming authority file from which the heading is taken. Use "lcsh" for LCSH, "aat" for AAT, or "gmgpc" for TGM I

http://www.cdlib.org/inside/diglib/guidelines/dmdreqs.html#dmdsubjecttopic
Centralized Vocabulary Work

- OCLC Terminologies Service
- High Level Thesaurus (HILT) project (targeted completion of phase 4 in early 2009)
- Simple Knowledge Organisation Systems (SKOS) leverages the Resource Description Framework (RDF) to support representation of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading systems and taxonomies.
Encoding controlled vocabularies for the Semantic Web using SKOS Core

Linda Cantara:

“Controlled vocabularies facilitate consistent documentation, but they do not guarantee flawless searching across multiple domains.”

we need for tools that can “link the relationships between search terms or phrases, tools that will facilitate searching for abstract and semantically significant content.”

Controlled Vocabularies & Subjects

• More transparency about vocabularies being used
• Work on formal cross walks between vocabularies
• Innovation into computer driven term grouping
What about user contributed tags?

What could the relationship be between EAD subjects and folksonomic tags?
Create Paths Where People Walk
‘Metadata for the Masses’, Peter Merholz

Photo from Phil Gyford(http://www.gyford.com)
Reprinted here under a Creative Commons Attribution-NonCommercial-NoDerivs 2.0
Cross Collection Understanding

INSTEAD OF

a text list of groups of records

RATHER

An understanding of quantity, subject and years covered by the combination of many collections
Information Visualization

- EAD opens the door to structured data about archival records
- Structured data lets us aggregate data – 30,000 foot perspective
- Standard used around the world
- XML is open and accessible
Challenges to Aggregation

• Diversity of subject terms
• Incomplete subject assignments – sometimes subject implied by title is not included as a subject
• Differences in standards for specifying dates and collection sizes
Future…

- Centralized Repositories and Union Catalogs
- Metadata Guidelines
- Thesaurii and subject cross walks
- Establish best practices for incorporating user generated tags in predictable and regular manner
- Information Visualization to give us the 30,000 foot view of collections
ArchivesZ

• Information Visualization class project
• Focus on EAD as a source of structured data
• Pull data into database for scrubbing
• Build UI based on aggregation and visualization
The Brooke Family was a large family of landowners in Maryland whose records relate to farm activities and family life on the plantation "Falling Green." The Brooke and Farquhar families were active members of the Quaker community in Sandy Spring, Montgomery County, Maryland. The collection contains the...