

Report from the Pilot EAD Implementation Project

Summary of the Project

As part of the LSTA grant, “Enhanced Access to Special Collections in the Libraries of Florida’s State Universities,” a pilot project was undertaken to build a union database of finding aids to archival and manuscript collections. Four institutions participated in this pilot: Florida International University, Florida State University, the University of Florida, and the University of South Florida. The University of Florida had one representative on the project team; the others each had two representatives.

Each institution selected three existing finding aids to encode in EAD format. Team members attended a training class given by the Society of American Archivists in the Encoded Archival Description (EAD), a standard for representing finding aids in SGML and/or XML. Copies of XMetaL, an SGML/XML authoring tool, were purchased with grant funds and distributed to the participants. Because the EAD standard allows a great deal of flexibility in how finding aids are encoded, an initial set of encoding guidelines was drafted to ensure some consistency between institutions. Each institution then encoded three finding aids in EAD format and submitted these to FCLA for loading into the union database.

FCLA used the University of Michigan’s DLXS software known as “Finding Aid Class” to make the finding aids available for end-user searching and display. This system is available at <http://purl.fcla.edu/fcla/findaid-test>. For the pilot period, the system was implemented “out of the box.” (Some modifications will be made in the future in response to usability testing.)

The terms of the grant required an end-of-project report evaluating the pilot and giving information on costs, problems, and benefits associated with EAD creation. Participants met in Gainesville in the last month of the project, September 2002, to compare their experiences and to outline this report.

Costs and cost factors

The amount of time spent on creating and editing EADs was recorded at each institution. These times are:

Institution	Total hours	Hours per EAD
FIU	101	33.5
FSU	91	30
UF	63	21
USF	34	11

This can be translated into costs, but the resulting figures would be extremely misleading, as the pilot situation was artificial in many ways. Some factors inflating the costs of the pilot include:

- The individuals participating in the pilot were in many cases a higher level of staff than staff who would actually work on EAD creation in a production situation.
- There was a long time lag between the SAA training and the start of EAD creation, leading to a steep re-learning curve. Also, pilot participants were the only ones in their departments with training in XMetaL, and had nobody to rely upon for support.
- Finding aids were deliberately chosen to represent a variety of cases, so each document raised new problems. Some were chosen because they were particularly long and troublesome.
- This was done simultaneously with another part of the grant project, the cataloging of uncataloged special collections. It was difficult to separately account for time spent on the two parts of the project.

Different methodologies were used at different institutions. Finding aid information had to be entered into an XML template in XMetaL. UF cut-and-pasted from existing finding aids in Word format. FIU first did optical character recognition (OCR) on paper finding aids, then cut-and-pasted from the OCR output. FSU tried OCR/cut-and-paste on one paper finding aid and found that simply keying it was faster. USF and FSU both had their holdings represented in relational databases, and were able to retrieve and reformat that information for the container list portion of the EAD.

Despite these differences some commonalties emerged that are likely to be true in a production environment:

Cost (effort) could be divided into two categories: verifying or improving the content of the finding aid, and encoding (entering data with XML tagging). In the first category the bulk of the work was to verify that finding aids accurately reflect the physical arrangement of the collection and that narrative portions of the finding aids are complete. In the second category the bulk of the work was in encoding container lists.

All participants found they had to do significant work to clean up the source finding aid before converting the information to EAD format. No participant felt it was acceptable to simply convert existing information to EAD format without making corrections, owing to the increased visibility of the online finding aid. No institution spent less than 25 hours on clean up work for their three finding aids; some spent considerably more. In many cases, this was the first time the finding aid had been verified since it was created. Old practices may have created less than optimal documents and EAD conversion is viewed as the time to correct past mistakes and improve on past practices.

Time spent on data entry ranged from a high of 71 hours to a low of 16 hours. Institutions who had already entered item inventories into a database spent significantly less time on this task. The effort of getting the information into the database to begin with was not counted as part of this project.

It should be noted that this project did not yield any data about what it would cost to create an EAD from scratch as part of the archival processing of a new collection.

The Union Database of Finding Aids

The online system for finding aids, called the “Florida Finding Aids Collection” during the project, allowed users to search the database and to display retrieved results in four formats: a collection summary, a keyword in context view, an “outline” view in which major section headings served as a table of contents, and as simple full text.

Usability testing was conducted with two FCLA staff and two library staff at the University of Florida. Three of the four testers had never seen a finding aid before. The testing resulted in a list of improvements which will be made to the system. Overall, however, response was uniformly positive. All of the testers liked the system and found it easy to use. The testers also enjoyed reading the finding aids themselves.

Conclusions

All participants thought the project was instructive, that EAD creation was worthwhile, and that an SUS-wide union database of online finding aids was a worthwhile goal. One participant noted that the project created “positive energy” toward Special Collections at her institution.

The online system for finding aids should be continued and moved to production status. As “finding aids” is not a widely known term, the system will be renamed “Florida Archival Collections.” Enhancements identified in usability testing will be made.

The encoding guidelines were helpful; no major changes are required at this time.

No institution has plans to retrospectively convert more finding aids to EAD format. Resources do not exist to do this. If resources did exist, some institutions thought that they would outsource the conversion and others thought they would do it in-house because of the amount of content work required.

FIU and FSU plan to create EADs for collections processed in the future. When FIU completes a series from their Office of the President there will be “critical mass” in the system for a public launch of Florida Archival Collections at FIU. UF and USF would like to create EADs prospectively but do not have the infrastructure or resources. Their continued participation, and the participation of other SUS institutions that did not participate in the pilot, may depend on whether a continuation project is funded.

All participants agreed the ideal situation would be to integrate the creation of finding aids in EAD format with the archival processing of new collections. The two major barriers to this are lack of staff and lack of training. Participants agreed that any staff working on EAD narratives would require training similar to that given by the SAA. In the ideal situation, a technical services professional trained in archives and manuscripts processing could create both the collection level catalog record and supervise the creation or conversion of the EAD finding aid. Paraprofessionals and graduate students could be trained in XMetaL, basic archival processing, and the EAD format. Student workers could be trained to enter item level information.

The project reinforced awareness that SUS special collections need specialized technical services support including professionals that understand archival theory as it relates to researching, arranging and describing collections. Participants agreed closer integration of traditional library cataloging and archival description would be beneficial.

The possibility of a statewide collection of searchable online finding aids is transformative. It has serious implications for collection development and collection support. Since collections would be far more visible to faculty at all state universities the acquisition and processing of collections would need to be seen in a statewide light. There should be some infrastructure to support cooperative collection building, especially for Florida materials. The broader agenda includes the development of special collections programs at individual repositories, the widening of customer service options for more developed programs, and consistent technical service support for underutilized and unrecognized special collections.

Future Action

The participants recommend that a grant be sought to build on the work of the EAD pilot project. There should be four main components to a continuation project:

- Develop a training program, similar to that offered by the SAA, but also covering the fundamentals of archival processing. The goal of the training would be to integrate the creation of EAD finding aids into the workflow of archival processing. Ideally an expert within the SUS would be funded to develop a curriculum and to visit the SUS universities to deliver the training on-site, using local materials as examples.
- Develop a manual for describing archival collections for use statewide. This could be a by-product of the training program.
- Develop a software application with a Web interface to support the creation and maintenance of EAD information in a relational database. At a minimum, it should support container list information, though the possibility of supporting the entire EAD should be explored. This should decrease the overall cost of EAD creation.

- Establish the infrastructure to aid in the development of a documentary strategy for Florida. Curators need to meet together, not only to develop a plan for cooperative collection development, but also to develop a persuasive rationale to help in soliciting both funding and donations of collections.

The Special Collections Task Force (SCTF) would be most appropriately positioned to work with FCLA on such a grant proposal. However the SCTF has no chair and has not been active for some time. The Digitization Services Planning Committee will ask the SUS library directors to clarify the standing of the SCTF.