OASIS ANNUAL REPORT, 2000-2001

Steering Committee: Kate Bowers, University Archives; Lisa DeCesare, Botany Libraries (secretary); Leslie Morris, Houghton Library (chair); Susan von Salis, Schlesinger Library; Julie Wetherill, OIS

Background: Work towards the OASIS system began with the Digital Finding Aids Project in February 1995. Harvard was one of a group of “early implementers” working to develop a national standard for SGML-encoded finding aids; version 1 of that standard, Encoded Archival Description (EAD), appeared in 1998. With the selection of a search engine (OpenText 5), and a Harvard-customized version of a web gateway that was originally developed by the University of Michigan's Digital Library Production Services, what is now OASIS became public in July 1998. HOLLIS provides a second route for access; OASIS requires that each collection represented by a finding aid also have a HOLLIS record linking to that finding aid.

OASIS began with seven participating repositories; membership has increased to eleven. Extensive information on OASIS can be found on the support site: http://hul.harvard.edu/ois/systems/oasis/.

The system includes 347 finding aids as of 25 June 2001.

Developments in 2000-2001: This year saw the migration of many functions formerly undertaken by OASIS participants to OIS, and the first round of system enhancements since the debut of the system in 1998. These can be summarized under the following headings:

–The OASIS support site. This web site, encompassing information on how to participate in the system, standards documents, historical material, training and support information, and “tips and tricks” with authoring software, opened in 1997. Material was written by OASIS members, and the site was designed and maintained by Lisa DeCesare, hosted on an OIS server. This year, responsibility for maintaining the site was transferred to OIS, and it has been completely redesigned to fit the OIS site “look.”

–EAD Workshops. Training workshops, taught by Susan von Salis and Kim Brookes of the Radcliffe Institute, have been offered six times since 1998. Coordination, previously handled by Leslie Morris, was handed off to OIS this year. To date, 90 individuals have been trained in EAD through these workshops. (The Workshop is a prerequisite for OASIS participation.)

–Standards development. Harvard has been an active participant in the development of the EAD standard, and this year submitted a number of suggested changes to the Society of American Archivists’s EAD Working Group. Harvard’s own encoding guidelines, which define local “best practice,” were first developed in 1996, and this year again have been expanded and refined.
System enhancements

- **Migration to XML/XSL.** Currently, finding aids in OASIS are encoded in SGML and delivered in both HTML and SGML. Both delivery mechanisms have drawbacks: the HTML offers limited navigation within the finding aid, a problem with long documents; SGML requires the user to download a viewer application, Panorama. Transforming the finding aids to XML and using an XSL style sheet to produce HTML combines the strengths of both. Since April, the OASIS Interface Committee has been meeting every other week to work with Lee Mandell on the design; the interim results can be previewed at: http://ois.harvard.edu/~lee/oasis/.

- **Linking to images.** Linking from the finding aid description of a letter or drawing to an image of the item was added to the SGML display. As part of the migration to XML, this capability will be extended to the HTML interface.

Steering Committee priorities for the coming year:

- **User survey.** The Steering Committee has had preliminary discussions with Prof. Terry Tivnan of the Education School about surveying techniques. Once the new style sheet is in place, a User Survey Working Group will be formed to gather feedback from archivists, reference librarians, and student/faculty OASIS users; this may well generate additional system enhancement requests.

- **Retrospective conversion.** LDI Access Grants have funded two projects, at Houghton and Schlesinger, for conversion of “legacy” finding aids. The Steering Committee will be monitoring both projects as part of ongoing planning for conversion of all Harvard finding aids. The significant increased number of finding aids in the system will also test the functionality of the OASIS’s search, retrieval, and display capabilities.
ENHANCEMENTS LIST

I. User-requested enhancements

1. XSL support for document delivery (UNDERWAY)
Replace current HTML delivery of finding aids with XSL-based delivery. Requires finding aids to be in XML rather than SGML (see #2). Lee Mandell has developed an XSL style sheet to start this process, but the finaid-idx.cgi script will need extensive changes.

2. XML support for finding aids (UNDERWAY)
OASIS should accept finding aids in either SGML or XML format equally, and convert them into XML for internal storage and indexing. This is needed both because contributors are beginning to use XML editors and because a migration towards XSL-based display requires finding aids to be in XML. Some issues that need research as part of this include: character set support mechanisms (i.e. unicode), and namespace definition for EAD.

3. Add support for intradocument linking (UNDERWAY)
EAD (and Panorama) supports links within a finding aid for doing things like linking controlled vocabulary index entries to corresponding sections of the document. For larger finding aids this is very valuable. The findaid-idx script should be changed to recognize these links and turn them into the appropriate elements in the HTML versions.

4. Navigation
Add to the navigation portion of the finding aid display an indication of where the user is in the finding aid. For example, a right-pointing arrow would indicate where the user is in the hierarchy on the left, as the user scrolls through the text of the finding aid on the right.

5. Questions or Comments (DONE)
The help links in OASIS to 'dfap@hulmail.harvard.edu' are sent to the entire OASIS user group. This was intended for help using the system, but is being used instead as a general reference service. Change link to go to libref e-mail account, with identification that query came through OASIS.

6. Support "archival master" finding aids in DRS
The OASIS contributors would like to be able to leave their master copy finding aids in the DRS, rather than store and track them locally. Add support for this to DRS so they can "check in and out" their finding aids as needed.

7. Character set support
OASIS currently supports the Latin 1 character set, and it needs to also support the ALA character set (per OIS policy decision in spring of 1999). An initial analysis is needed to define the set of characters in ALA that are NOT in Latin 1. There are problems of input, indexing with OT, and online display. Related to item #2 (implementation of unicode may resolve this problem).
8. Browse list navigation
Add alphabetic index to top of browse list to aid navigation in very long lists of finding aids (e.g. Houghton). Should be produced automatically during construction of browse lists by Update_Finding_Aids script.

9. Browse list display
Add file size and version date to finding aid browse lists generated by findaid-idx.cgi.

10. Large document delivery
Houghton is producing finding aids of up to 4Mb in size. Much too large to deliver in HTML as a single file to the browser. We need to consider how to accommodate very large finding aids in the interface. Requires analysis -- not a problem yet.

11. Enhanced security
University Archives would like to be able to restrict access to portions of some of their finding aids entirely (i.e. not even accessible to Harvard affiliates). The example given was administrative documents which include a series on "Disciplinary Cases" which have folders titles of individuals under disciplinary review.

12. HTML conversion time
The findaid-idx.cgi and the preview.cgi scripts both do EAD->HTML conversion on-the-fly. This script needs considerable cleanup to make it faster, or else we need a new architecture that allows for pre-conversion of the SGML documents to HTML (involves removing lots of stuff from findaid-idx.cgi -- very tricky). Requires analysis -- not really a problem now.

II. OIS-required enhancements

1. Move-to-prod
High priority
Add a scripted "Move-To-Prod" facility for program and HTML page changes using "snap" and following the HOLLIS Plus model as applicable.

2. Update_Finding_Aids script fix (DONE)
High priority
Change Update_Finding_Aids to reindex the site with a switch instead of requiring there to be a finding aid in the ftp drop box.

3. Replace OpenText search engine with alternate XML search engine
Medium priority
OpenText is no longer supported, except in a modified form by the U of Michigan as part of its DLXS program.