THE OHIOLINK EAD FACTORY: CONSORTIAL CREATION AND DELIVERY OF EAD

CARA GILGENBACH, AMY MCCRORY, AND DAVID GAJ

ABSTRACT: This article describes the development and implementation of the OhioLINK EAD Finding Aid Creation Tool and Online Repository project. Collectively referred to as the OhioLINK EAD FACTORy, the consortial project includes two major resources—a Web-based application that allows for the creation of EAD finding aids without the user needing to know about EAD tagging or code and a centralized finding aid repository that allows EAD from across the state of Ohio to be searched and browsed. The authors include a discussion of how the project, led by an OhioLINK-appointed task force, evolved, objectives of the project, and outcomes to date, as well as issues and challenges encountered. They also detail the functionality of both the encoding tool and repository search engine, training programs for users, and future needs and goals of the project.

Introduction

The creation of a consortial environment in Ohio for encoding, displaying, and searching archival finding aids has taken place against the backdrop of many years of collaboration among libraries. The state of Ohio has been recognized for its development and implementation of cooperative library programs, most notably including the Online Computer Library Center (OCLC), originally established as the Ohio College Library Center in 1967. Twenty years later, the state created the Ohio Library and Information Network (OhioLINK), a cooperative academic library system that established joint purchasing of electronic resources and provision of interinstitutional lending across Ohio. Founded in 1987, OhioLINK is a statewide library consortium, now comprised of 88 members, that brings an extensive array of resources and services together to meet the information and research needs of Ohio students, faculty, staff, and a broader community of citizens. "OhioLINK and its member libraries provide access to 46 million books and other library materials, millions of electronic articles,

12,000 electronic journals, 140 electronic research databases, and 25,000 e-books," among other resources.[1]

OhioLINK has also initiated projects that support institutions in making their own unique content and local collections more widely accessible to both state-level and worldwide communities of users. In 1999, it established the Digital Media Center (currently evolving into the Digital Resource Commons[2]), which archives and provides centralized access to both commercially purchased multimedia content (such as educational films and art images) and unique collections of materials digitized and contributed by member institutions (such as archival photograph collections and audio recordings of oral histories).[3]

One of OhioLINK's operational committees, the Database Management and Standards Committee (DMSC), includes members who primarily specialize in technical services, cataloging, metadata, and systems operations. The DMSC had taken a leadership role in producing metadata standards for the OhioLINK Digital Media Center, but was not in a position to oversee development of standards for specific encoding schemes, such as Encoded Archival Description (EAD). It was generally ascertained that many institutions, though interested in adopting EAD, lacked the expertise, systems support, and staffing to easily or effectively utilize it.

EAD is one of many varieties of eXtensible Markup Language (XML). XML is a language that allows for meaningful structuring of document data, as well as formatted presentation of the data for electronic transmission, primarily through the World Wide Web. EAD is a type of XML designed for the specific needs of archivists as they describe the collections in their holdings. It allows for structuring of traditional finding aids in a standardized set of data fields, making it possible to index, search, and display their contents online. EAD was created as a Document Type Definition (DTD), essentially a set of instructions about how the document may be structured.[4] Although EAD has been accepted as an international standard, many practitioners have found it difficult to effectively adopt and administer it.[5]

In an effort to encourage adoption of EAD and metadata standards for digital archives, a number of consortial and state-based projects, such as the California Digital Library, North Carolina's NC ECHO program, and the Kentuckiana Digital Library, have arisen. Despite differences in the specifics of implementation, these and similar projects across the country tend to include three major components: use of content guidelines or standards for preparation of finding aids and metadata for digital objects, provision of templates or other tools to assist in the EAD encoding of finding aids, and consortial delivery and centralized searching of finding aids and other digital files to enhance access to archival collections. The OhioLINK EAD project articulated many of these same goals, primarily the creation of content standards for all EAD finding aids produced and the development or utilization of an encoding template or tool that would make EAD adoption as easy as possible within the state of Ohio.

In 2002 and 2003, The Ohio State University Libraries (OSUL, an OhioLINK member) was developing templates for encoding finding aids in accordance with the EAD standard, with the goal of producing a system that could be used across all its special collections departments. Finding aids encoded in EAD were converted to HTML and uploaded to OSUL's Special Collections Web site. However, OSUL Special Collections

staff also wanted data search and retrieval capability for these documents. This would have required implementation of complex systems software, which could not be accomplished in-house with the resources available at the time. Accordingly, they began a dialogue with OhioLINK staff about consortial delivery of EAD that would benefit institutions statewide. OhioLINK, in turn, was interested in developing such a project, but stated the need for a group of people who could guide the process and, as EAD is a specialized metadata standard tailored to archival practice, write clear and specific content guidelines. The OhioLINK DMSC appointed an EAD Task Force in late 2004, with the general charge of finding a way to implement the creation of EAD finding aids within the OhioLINK community. A more detailed charge given to this new task force included the following directives.

1. Develop a specifications template (content standards) for EAD documents created within the OhioLINK community.
2. Follow the progress of EAD software enhancement requests (at the national level).
3. Customize a software toolkit for the creation of EAD finding aids in a multi-institutional environment.
4. Generate marketing and training ideas to encourage the use of EAD among OhioLINK members.
5. Educate members of the DMSC on progress and findings.

The Task Force's roster, selected by DMSC members, was comprised of archivists, special collections librarians, catalogers, and metadata specialists.[6] Some Task Force members had strong technical backgrounds in EAD, metadata, and cataloging, while others had less technical expertise but had significant experience in providing access to collections, working with researchers, and responding to requests from off-site users. This proved to be a good mix of skills and sensibilities when it came time to begin working on EAD content standards, one of the group's first major projects.[7] For example, those with technical experience with EAD were able to identify which elements would need to be included for a valid finding aid, and were also familiar with common encoding practices, while those with stronger public service backgrounds were familiar with the type of information about collections researchers commonly seek and which fields might be required within the OhioLINK setting.[8] The Task Force also periodically consulted with selected personnel from historical societies, both large and small, in an effort to help ensure that the content guidelines and its general approach to the project were realistic and on target from the perspective of non-collegiate institutions.

Inherent in this broad appeal to practitioners in diverse settings was the desire to make EAD creation in Ohio a possibility for all types of institutions. The specifications template would have to satisfy the requirements of archivists serving different types of users, and therefore institutions would have very different, locally driven descriptive priorities. Fortunately, EAD makes possible this sort of diverse practice. However, the standard's very flexibility can lead to problems in creating a collaborative discovery environment, one in which an XML-aware search engine is programmed to comb a large collection of documents for regularly encoded pieces of data, and display them uniformly.[9] Task Force members decided that, in the interest of standardization, the OhioLINK EAD project would restrict contributors to certain patterns of encoding,

while allowing for as much variation in finding aid creation as possible. For instance, many institutions consider it valuable to include highly granular encoding of personal names in association with individual items in archival collections. The Task Force decided not to make this capacity a part of its encoding template or tool, however. Offering the option of including the large set of EAD elements available, in various configurations, at increasingly specific levels of description would make guidelines for encoding practice overly complicated, working against the goal of making the guidelines easily understandable for people at all levels of technical expertise. In addition, it would work against consortial delivery of finding aids through a shared search engine, as noted above.

Discussions of standards for the project offered surprises from time to time. One group of Task Force members assumed a biography or history note for each collection would be required. Another group took issue with this, arguing that in the real world of archival practice, the personal or corporate history of a collection's creator may be unavailable or collections may be compiled by an institution from numerous sources. Protracted discussion took place over this issue, until it was decided that the biography/history element would be designated "mandatory only if applicable." The discussion occupied more time than anyone on the Task Force had expected, but it became a good example of how dialogue between the technical practitioners and the managers of archives was useful in drafting standards that accommodated the day-to-day realities of archival practice.

The group's agreed-upon goal, then, became to write OhioLINK EAD content standards that took into consideration what the majority of finding aid authors would reasonably want to include in their encoding practice. Development of these standards took two years to complete. This effort, though at times quite arduous, proved to be of great benefit to the development of an encoding tool and online centralized finding aid repository, both of which were created during the second stage of the OhioLINK EAD project.

Encoding Tool and Finding Aid Repository

Encoding Tool

During the initial phase of exploration and development of finding aid content standards, Task Force members had also explored possible options for the development of a template that could be used by encoders in any standard XML editor. This notion, however, eventually gave way to the idea of creating a Web-based application that would offer a number of advantages to its users.[10] It would be accessible (with a login) from any computer with Internet connectivity and would not require the download of any piece of software, other than a Web browser. Ideally, the application would not require its users to have any technical knowledge of EAD XML tagging and would make all tagging completely invisible to the encoder. A wide range of finding aids, from simple collection-level descriptions to more complex and detailed documents including component-level description, would be accommodated.[11] The application would allow

for copying and pasting of information from legacy documents into its interface. It would also offer the ability to locally download the raw XML file produced (for local archiving and other purposes). Additionally, it would facilitate uploading of any EAD finding aid created outside of the encoding tool (legacy EAD previously authored at local institutions, for instance), so long as the EAD was valid. Finally, user account creation and management would be as simple as possible, and would be controlled locally at each institution that registered to use the tool.

Perhaps the most significant objective for the application was that it not require users to know any XML EAD tagging. Task Force members realized that the technical complexity of XML tagging and the overwhelming number of EAD tags and attributes were often the biggest roadblocks to adoption of EAD. Explorations of XML editor software, even those utilizing templates, reinforced the notion that dealing with tagging on any level was likely to hinder widespread use of EAD. Furthermore, a review of professional literature on EAD adoption and reports of other consortial projects bore out the notion that one of the most effective ways to encourage adoption of EAD was to make encoding and publishing of finding aids to the Web as easy as possible. In a recent article on barriers to implementation of EAD, Sonia Yaco discussed the results of her survey of a range of institutions across the country. She concluded that there are three major impediments to adoption of EAD: lack of staff, the so-called middleware gap (an inability to deal with the technicalities of mounting EAD on the Web due, in part, to the lack of systems support), and the need to revise finding aids extensively before encoding.[12]

In its discussions with potential users at various presentations and meetings, the Task Force heard similar challenges voiced in relation to EAD adoption.[13] For instance, many institutions indicated a heavy reliance on student and volunteer staff to create finding aids, and a lack of time or resources to train such staff in the technicalities of EAD. In many cases, permanent or full-time staff also lacked adequate training in EAD encoding. Others not only lacked the expertise to create EAD finding aids, but they also had created no Web-mounted finding aids to date and were considering ways to present even simple, collection-level finding aids on the Internet. Even staff members who had been able to attend some EAD training expressed the limited success they had had with actually administering EAD at their institutions, often due to the lack of administrative and systems support in particular. Thus the Task Force determined that creating a tool that could address these issues would be better than simply offering templates for use in XML encoding software.

An individual Task Force member forged a connection with the Kent State University Libraries (KSUL) Systems unit to explore the possibility of codevelopment of such an application. The KSUL Systems unit has, for many years, developed customized Web-based applications to address a variety of needs, and the group had been in the tentative stages of planning for a tool that would allow the Kent State Department of Special Collections and Archives to create EAD finding aids on its own. If local development of such a tool was likely to take place, why not expand this development to the benefit of other OhioLINK institutions, was the thought that led to the eventual partnership with the EAD Task Force.

Once the commitment to this partnership was made, the encoding application was created during an intensive period of development from November 2006 through September 2007. The first order of business was for Task Force members to provide developers at Kent State University with specifications for the application, including all of the required and optional fields and their corresponding tagging structure, field labels, selection list values, and tagging specifications for the XML code that would be generated by the application. At the same time, the Task Force had to envision how the application should function from a user's perspective, what menu items should exist, and how an encoder should best work through the various stages of finding aid creation. There also needed to be a mechanism in place for storing data and allowing for files to be uploaded to a centralized finding aid repository (which was also being planned, but had not yet been created). Decisions had to be made about what elements or fields should be allowed at both the collection and component levels of description. Specifications also needed to address date normalization, embedding of ISO language codes,[14] and choices regarding MARC record encoding analogs[15]—and, above all, how to present these options to users with as little technical jargon as possible.

Some of the major issues encountered in the application development process included choosing what subset of EAD fields the application would support, how to implement formatting elements such as the various list styles allowed in EAD, and how best to allow component-level encoding. Another key concern was how the finding aids produced in the application would be uploaded and displayed in the public, search engine/repository environment. Additionally, decisions made during the creation of the application led to a reassessment of the content guidelines written during the initial phase of the Task Force's work. The original guidelines contained a good deal of technical detail, with encoded examples of each element. In light of the application's development as a product accessible to archivists without extensive EAD training, content guidelines were modified to emphasize the content and format of each element, not the XML tag set and its position in the encoded hierarchy. In their new form, the content guidelines took on a second function as a set of contextual help pages accessible from within the encoding tool.

Technical Architecture and Components of the Finding Aid Creation Tool

The OhioLINK EAD Finding Aid Creation Tool is comprised of a Web-based application written primarily in PHP (an open-source scripting language geared toward writing Web-based applications) and HTML.[16] JavaScript is also used to support dynamic generation of fields and screen-based validation. Finding aid data is stored in various tables within a MySQL database (an open source database system used widely in Web-based applications). For example, there are MySQL database tables named "accessrestrict" and "relatedmaterial," which correspond to those specific elements of the EAD. Nested component-level information is handled by storing the structure of the entire component level in the database, complete with the "parent" and "position"

of each component level within the hierarchy. This highly structured database lends itself easily to rendering an EAD-compliant XML document.

In order to register to use the application, the requesting institution completes a publicly accessible form on the Web, which is submitted to an EAD Administrator, who oversees all institutional accounts within the application. The EAD Administrator reviews the information and creates the institution's profile within the application. This process automatically creates an administrative account for that institution and generates an E-mail to the administrative account holder containing his or her login information and simple directions for setting up local user accounts. The administrative account holder then has the power to create and edit all of the local user accounts needed at his or her institution. Access to the application is only granted to those with valid accounts; each account is configured to limit user access to the appropriate finding aids, depending on the access level assigned to each user.[17]

The user enters his or her finding aid information into the application via Web-based screens that are separated into the various components of an EAD finding aid such as the EAD Header, Collection Summary, Scope and Content, Preferred Citation, and a number of optional fields. Screen-based validation helps to catch errors upon entry, guiding the user toward an EAD-compliant resultant finding aid.

Although there are several additional functions that can be performed at any time (by clicking the appropriate link), they are most useful once a finding aid has been completely entered. These include validating the finding aid, previewing the finding aid, downloading the finding aid to one's local computer, and submitting the finding aid to the public repository (although uploading is not offered as an option until the finding aid has been properly validated). Behind the scenes, an XML file is automatically generated, which facilitates the operation of these functions. Validation of the finding aid occurs by specific programmatic checks being performed against the data, followed by verification of the previously generated XML file against the EAD DTD. Previewing the finding aid content is accomplished by transforming the previously generated XML document into HTML via an eXtensible Stylesheet Language (XSL) stylesheet.[18] Downloading a finding aid to the user's local computer is performed by prompting the user to save the previously generated XML document to his or her computer. Finally, submission of a finding aid to the OhioLINK Online Repository (discussed below) is executed by copying the previously generated and validated XML file to a "staging" area on the server. An automated script runs nightly to copy all submitted XML files to the public repository server.

The application's user interface consists of a series of Web-based forms, some of which are required (denoted by italicized, starred field labels), and others that are optional. The application utilizes standardized graphics throughout to designate various functions such as clicking on a pencil icon to enter into editing mode for a given field, or clicking on a question-mark icon to view the help page for a field. Rollover text is included with these icons in case one does not recall or initially know what the icons mean. The user (encoder) can opt to display all possible fields available in the left menu bar of the application or can hide fields that are optional. This feature can be helpful to the inexperienced user who may feel overwhelmed by the number of optional field choices displayed in the menu area. Some fields utilize drop-down menus that allow

the user to select his or her choice from a range of standard values. For example, in the Extent field, a user can choose from "cubic feet," "linear feet," or "boxes" via a drop-down menu. Additionally, the application features what Task Force members refer to as "validation on the fly" in that a given section being edited cannot be saved unless required data fields are completed. This type of behavior is encountered on other commonly accessed Web pages such as on commercial sites that utilize forms for placing orders or survey forms utilized on a wide range of sites. Validation of this kind helps to ensure that required information is filled in as one works through the finding aid, instead of having to wait until a final validation check to locate instances of uncompleted required fields.

While a user must perform a validation check of the finding aid before it may be submitted to the Repository, he or she can utilize the application's validation function at any time when creating the finding aid. This can prove to be a useful exercise, particularly for beginning encoders, who will see which required areas remain to be completed or will be alerted to other validation issues as the finding aid is being constructed. There is also a finding aid preview function that allows the user to generate a preview of the finding aid at any time during encoding. This can help the user maintain a sense of the finding aid in its entirety, as data is being entered into the various sections of the application interface. The preview feature also allows the user to print out a copy of the finding aid for proofing or editing purposes, before it is downloaded locally and submitted to the OhioLINK Repository.

In addition to setting up local user accounts, administrative account holders have the ability to change institution profile data, including information such as institution name, address, and Web page URL. When information in these areas of the profile is changed, any finding aids with a status of "submitted" (to the Online Repository) are automatically resubmitted to the Repository by the application, so that the changes will be applied to all of that institution's published finding aids. Administrative account holders can also set institutional preferences for default language for some fields, such as Preferred Citation. This saves input time for users who do not have to retype standard language used in selected fields in each and every finding aid. It also allows the institution to maintain consistent language for selected fields. For a small number of fields, the application also offers the option to suppress public display of data entered, if the institution so chooses. These are fields that may contain information that institutions want to keep confidential in the Online Repository, such as physical location of the collection.[19]

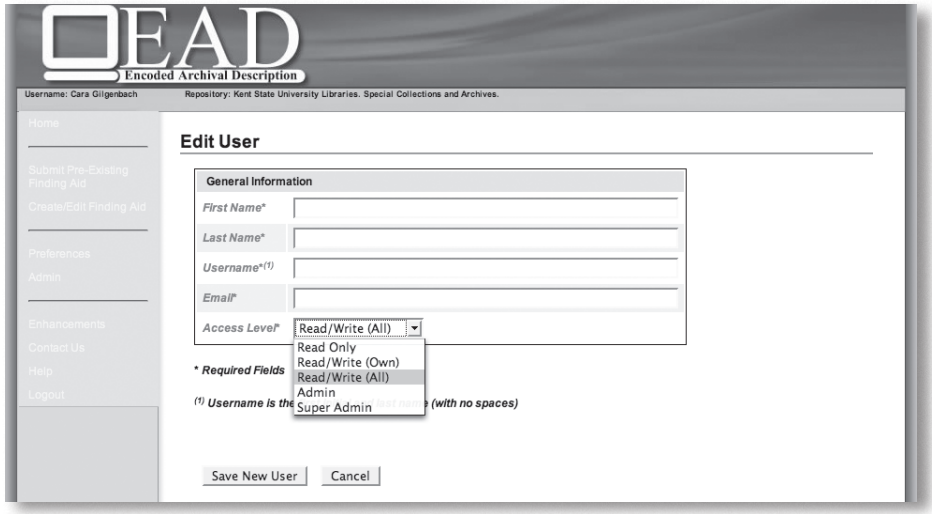

Illustration 1. The Finding Aid Creation Tool interface for creating a new user account.

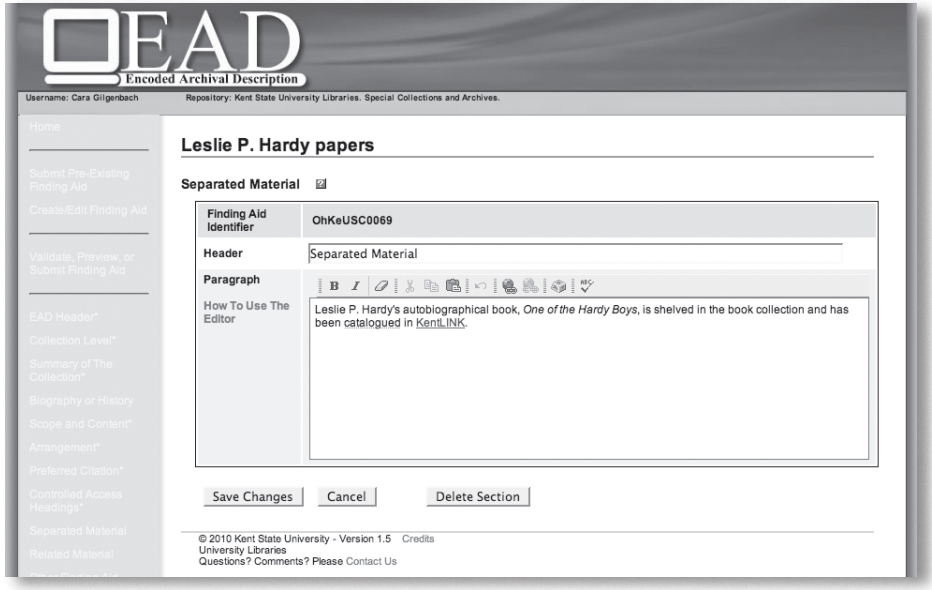

Illustration 2. A sample collection-level input screen in editing mode. Note that menu options to the left are made inactive (ghosted) until the editing screen is saved or canceled.

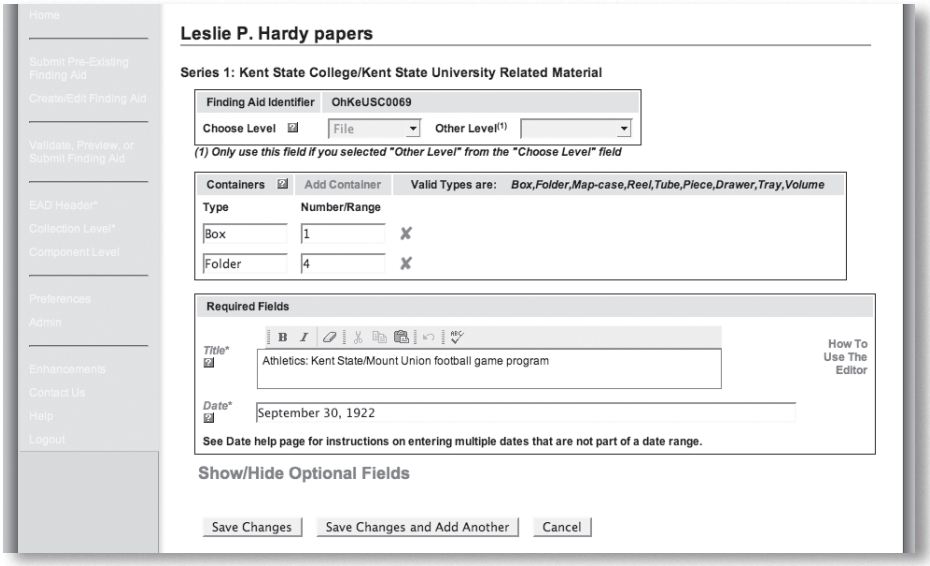

Illustration 3. A component-level input screen for a file-level item being described. There are a number of optional fields that can be attached to any component-level entry by clicking on the "Show/Hide Optional Fields" link.

Finding Aid Repository and Search Engine

At the same time as the encoding application was under development, work was under way to create a centralized, public repository where finding aids from participating institutions could be displayed and searched by the public. During the initial phase of work from 2004 to 2006, OhioLINK programmers tested three freely available search engines designed to work with EAD, including XTF, an open source software package developed by the California Digital Library (CDL). In 2006, OhioLINK chose XTF for delivery of its TEI (Text Encoding Initiative) E-books project. Given the similarities between the EAD and TEI standards, and the staff expertise in XTF development, OhioLINK decided that XTF should be adopted for EAD as well. Configuration of XTF (version 1.9) for the finding aid repository was undertaken and continued throughout 2007. The Task Force and OhioLINK systems developers were fortunate to have the cooperation and support of the XTF developers at the CDL, who provided advice early in the process, and included OhioLINK's senior systems developer in the testing group for new releases and upgrades of the software.

XTF, to put it very simply, consists of two layers of code, one of which is familiar to anyone who has modified standard XSL for EAD, the other of which is particular to the way XTF indexes and delivers data. There are two important points to be taken from this. For the first layer, one can simply modify the existing style sheet in order to configure the output to local customs or preferences. For the second layer, it is necessary to have a programmer who is familiar with XTF and can adapt the changes in the first layer to work with the second layer. The Task Force was fortunate to be able to work with OhioLINK's XTF programmer, whose expertise made it possible

to configure the search and display features of the software according to the group's recommendations.[20]

Concurrent with the development of the software, the Task Force began making recommendations for a gateway page consisting of a simple search interface; a link to an "advanced search" page; a link for content contributors, connecting them to a page explaining how to sign up to use the EAD encoding tool; and options for browsing, rather than searching, the collection of finding aids. Since the Task Force was adapting OhioLINK's E-books application of XTF, it was necessary to redefine and rename fields and browsing options. This gave the Task Force the opportunity to focus its ideas about categories of accessible information, as well as the best terminology to employ in this interface designed for researchers and other end-users, that is, non-archivists.

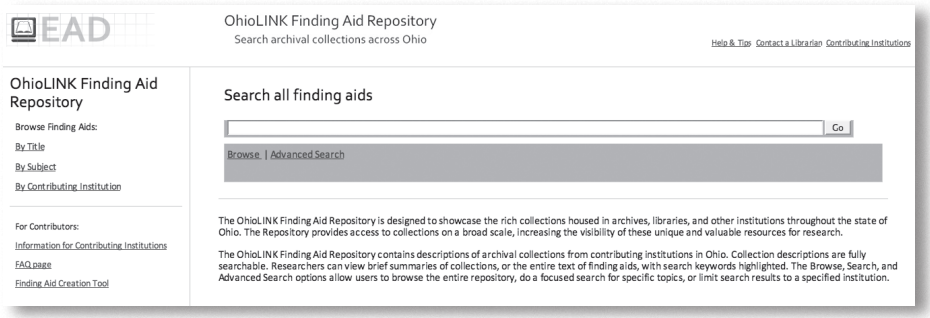

Illustration 4. The homepage for the OhioLINK Finding Aid Repository.

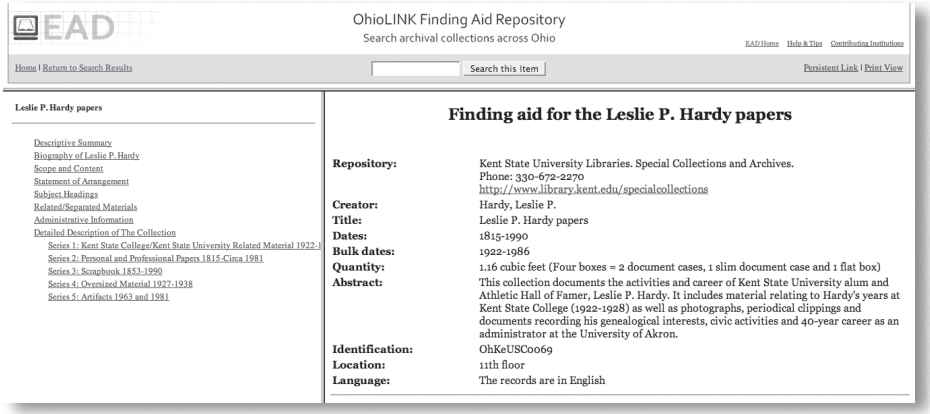

Illustration 5. A portion of a finding aid in default display mode. Note that the body of the finding aid is to the right, while navigational links to other portions of the file are available to the left. The "Print View" option allows the user to view the finding aid in its entirety in a printer-friendly format.

The EAD FACTORy in the Context of Other EAD Implementation Systems

The EAD FACTORy is one of many systems developed in the past decade to facilitate encoding of archival finding aids. Comparison to other such efforts helps to elucidate the particular decisions the Task Force made during the development of this project.

Two of the most significant EAD encoding products are the Archivists' Toolkit™ and Archon™. The Archivists' Toolkit™ offers a complete system for managing archival activities, from acquisition of a collection—with its attendant record-keeping requirements—to collection organization, and description in the form of a valid EAD finding aid. It is a powerful beginning-to-end tool for archives management, initially tested among a select group of universities and now used in institutions across the United States and in Europe.[21] Archon™ is a software package that makes it easy for an institution to create finding aids, MARC records, and digital object repositories for its local collections, integrating all three through linked provenance and subject headings, and generating a searchable Web site. As with the Archivists' Toolkit™, users across the United States and in Europe have downloaded Archon™ for use at their institutions.[22] While the Task Force recognized the value of both of these freely available systems, its members concluded that too many institutions in Ohio would decline to adopt either one, because both require that software be downloaded and installed locally. Even when the acquisition, installation, and use of software can be done without extensive local systems support, some archivists view this activity as too time-consuming, challenging, or demanding of scant local resources. Whether or not this perception is accurate, Task Force members agreed that it would be a barrier to statewide implementation of EAD. For much the same reason that distribution of XML encoding templates was ultimately set aside in favor of a Web-based interface, the Task Force decided to proceed with a new encoding tool and searchable repository, to be centrally hosted at Kent State Libraries and OhioLINK, respectively.

A pleasant surprise came, after the release and statewide adoption of the EAD FACTORy, when one institution reported using it for finding aid creation, while using Archivists' Toolkit™ for archival management and production functions. This suggests that the available tools supporting archival activity may be used in conjunction with one another, rather than exclusively. In fact, participation in the OhioLINK Online Repository does not *require* the use of the OhioLINK encoding application (although use of the application is strongly encouraged). Institutions with legacy EAD XML files and those that want to encode finding aids via another tool or application may still register with the project. In this scenario, the institutions utilize the encoding application only as a "pass through" mechanism to upload EAD files into the Repository. As long as numerous EAD tools are freely available, archivists will be able to decide how best to adapt them to their own needs.

Participant Feedback and Issues Encountered in the EAD FACTORy Project

The first public demonstration of both the encoding application and finding aid repository occurred in October 2007 at the Academic Library Association of Ohio conference, where the project was generally met with enthusiasm and interest. This

led directly to a period of beta testing of the application with a larger group of test encoders who, importantly, had not taken part in the tool's development and had not used it before. Following beta testing and completion of modifications arising from that testing, the OhioLINK EAD Finding Aid Creation Tool and Online Repository (referred to collectively as the OhioLINK EAD FACTORy) were officially released for production use in early April 2008. To date there are 39 institutions, including some public libraries and other non-collegiate institutions, using the Finding Aid Creation Tool and submitting finding aids to the Online Repository. Currently, 121 individual user accounts are registered in the system. Several institutions report having registered volunteer and student staff to use the application to create finding aids. One institution has utilized the Finding Aid Creation Tool in conjunction with a public history graduate course in which each student is tasked with processing and then creating an EAD finding aid for a given archival collection.

In general, users have reported that they find the Finding Aid Creation Tool easy to use. Staff at a handful of institutions have noted that this is the first application they have used to create electronic finding aids, indicating that existing finding aids have not been accessible online in any format. Others have noted that the application is accessible enough to allow them to assign EAD finding aid creation to volunteer and student staff members. An archivist at a small liberal arts college, who is the archives' lone staff member, offered the following feedback:

> As someone with just enough technological skill to be dangerous, I can say that the tool has been very helpful, and as improvements and tweaks are made to it I think it's getting even easier to use. It allows the user to decide just how detailed and/or complicated a finding aid you wish to create. However, I have learned the hard way to carefully assess my in-house finding aids better to see if I can use them as a source for the EAD finding aid.

This archivist also stated that while she still did not have a lot of time to devote to finding aid creation, she has included in her department's strategic plan a goal of producing three to four new EAD finding aids on an annual basis.

Archives staff at a midsize state-funded university provided feedback indicating that the "tipping point" in the decision to implement EAD came when a viable consortial project emerged.

> Like many institutions, the department had long discussed creating EAD finding aids, but as a lone institution, the advantages were limited. The time in coding finding aids in EAD seemed to yield little advantage when our finding aids could be searched on the Web using a site search. While not a precise search, the resulting hits were not overwhelming. However, as a contributing institution of a *cooperative* project like the OhioLINK EAD Finding Aid Repository, the advantages to participation are enormous.

Among the advantages cited by this institution is that the encoding application allows the generated XML to be easily downloaded locally. The option of automatically generating a basic MARC record, which can save considerable time for institutions

that produce such records for local OPACS and worldwide catalogs, was also viewed as a benefit. Finally, some of the advantages to end-users were noted.

> Our institution and research collections gain visibility from participating in a statewide repository as researchers search across collections in the state. Researchers are able to search our finding aids with greater precision by searching on meaningful fields while still having the flexibility to search by key word. We [the archives staff] are more knowledgeable about archival collections across the state related to our subject areas.

Additional informal interactions with contributing institution staff have generally yielded positive comments, along with enthusiastic suggestions on ways in which the encoding application and the finding aid repository might be enhanced. More formal recognition of the impact of the project came in 2009 when the project was awarded a Society of Ohio Archivists Merit Award, which recognizes "individuals or organizations that have by excellence in deeds, actions, or initiatives improved the state of archives in Ohio over the past year."[23]

Despite the generally positive feedback from participating institutions, there are challenges and issues that have arisen in relation to use of the encoding application. The two areas that consistently appear as more complicated are the templates for controlled access headings and the succeeding component levels (such as series, sub-series, file, and item) in the Detailed Description of the Collection section of the finding aid. No doubt this is due to the multilevel nesting inherent in these portions of EAD. Translating this structure to a simplified graphic interface was one of the more difficult parts of the application's development. Careful explanation appears to eliminate confusion in these areas, and so these sections of the Finding Aid Creation Tool are given ample time and emphasis in training workshops that are offered. The Finding Aid Creation Tool also allows the EAD Administrator to send out system-wide user messages that can address encoding problems as well as send out information on enhancements or changes made to the application.

While the Finding Aid Creation Tool provides a means of structuring content according to the rules in the EAD DTD, it does not exert any sort of control over the content entered in each defined section, with the exception of a few fields, such as those designating MARC values. This does, inevitably, allow for inappropriate use of data fields. For instance, the inclusion of container numbers or scope and content notes in a field meant solely for the encoding of unit titles has been noted in some finding aids produced. Indeed, this is a possibility that exists within EAD itself and could occur regardless of the software used to create the encoding.[24] The short-term consequence is the possibility that such finding aids will display incorrectly in the Online Repository. Longer-term implications include possible incompatibility with other consortial systems to which an institution might someday want to contribute its EAD files.[25]

The EAD Task Force has discussed how far to go in "policing" the use of data fields. To prevent such problems, use of the content guidelines and an "EAD Starter Kit" posted at the OhioLINK Web site, along with the contextual help pages included in the encoding tool, are heavily emphasized in training materials and when project leaders respond to requests for help by individual users. Ultimately, it is up to each contributing institution to study the guidelines and adopt the recommended best practices. A

clear-eyed approach to building a cooperative EAD environment includes understanding that while valid document structure can be enforced (through the validation process described above), it is not possible to exert complete control over content placement within that structure.

Issues encountered with the deployment of the Online Repository primarily have had to do with proper display of finding aids, along with some concerns over more effective navigation. An anticipated area of dissatisfaction was the display of finding aids. The development team was concerned that the unified style of presentation offered would fail to meet the needs of encoders from different institutions with diverse styles of presenting finding aids. There has been, however, very little complaint with the style chosen. Requests for changes to the presentation are evaluated on the basis of whether they can be successfully applied across all finding aids. The bulk of the enhancement requests received in the first release of the repository were documented for implementation during the upgrade of the software.

Project Support and Future Viability

Work on this project moved forward without a budget being produced up front. While this seems unusual, it is a somewhat common practice within the OhioLINK community setting. Institutions generally have "bought in" to the notion that work for OhioLINK on behalf of statewide users is part of normal work life and job responsibilities of staff at member institutions. For instance, systems development time was "donated" by the Kent State University Libraries to allow for the creation of the Finding Aid Creation Tool and by The Ohio State University Libraries for Finding Aid Repository development. OhioLINK has allowed a senior systems developer to work on the Repository side of the project as time and other, higher priority projects have permitted. Task Force members were granted time by their home institutions to attend meetings, write content guidelines, conduct training workshops, and do presentations at statewide conferences. In all cases, key staff were able to garner institutional support that allowed them to devote some work time to the project.

One of the greatest risks facing the project, however, is the loss of this institutional support in a time when nearly all state-supported agencies, and many private institutions, have implemented hiring freezes and restrictions on filling vacated positions. Staffing is a major concern statewide, and could affect ongoing development and maintenance of the initiative. The reality is that much of the work is sandwiched in between competing projects. Additionally, state budget cuts could affect the viability of the project if those cuts significantly reduced staff at OhioLINK, particularly if cuts involved key systems staff who contribute to the project. The EAD project is not a full-time concern of any staff member at OhioLINK or at the Task Force members' institutions.

The first four years of the project give reason to be optimistic, however. Two years into the Task Force's existence, half of its members took jobs or educational opportunities outside the state of Ohio, ending the participation of some of the project's key developers. The remaining members were able to recruit new people and successfully continue development of both the Finding Aid Creation Tool and the Online Repository, as well as designing a series of workshops and related training materials. It is expected that

institutions that have already created and contributed finding aids, as well as those that have devoted staff resources to the EAD FACTORy, have an interest in the long-term success of the endeavor and will continue to contribute staff time to make it possible.

Continued systems support is a critical issue, as a specialized skill set is necessary to develop and support the complex architecture of both components of the EAD FACTORy. Kent State's contribution of programmer time is a fine example of how local expertise can be donated to strengthen consortial projects. The Task Force is interested in garnering similar support from university libraries throughout Ohio, as a widely distributed system of donated staff hours would provide a safeguard against loss of personnel at a single institution.

Current and Future Objectives

The current initiatives of the Task Force include creation and implementation of in-person and online training programs and more targeted outreach to libraries, archives, historical societies, and museums throughout the state of Ohio. In the future, attention must also be given to end-user interactions with the Online Repository in order to gauge the effectiveness of this site for researchers. The Task Force is also engaged in enhancing and upgrading both the encoding application and Repository on a continuing basis.

Training

In 2008, the Task Force conducted its first in-person, hands-on training in the form of a daylong preconference workshop during the Society of Ohio Archivists annual conference. Subsequent workshops have been offered in various regions of the state. The workshops, designed and taught by two members of the Task Force, focus on giving attendees an initial orientation in the use of the encoding application and allowing ample time for participants to gain hands-on exposure to the application as well as the finding aid Repository. A two-hour morning session is devoted to learning about the collection-level input fields, while two hours in the afternoon are devoted to completion of the optional component-level areas of the finding aid, along with review of the final steps of validating, submitting, and locally downloading the finding aid produced. Each demonstration period (no longer than 25 minutes in length) is followed immediately by a hands-on session to allow participants to try out what has just been covered. The time allotted for the hands-on work is equivalent to the time for demonstration. Instructors are available during the hands-on sessions to answer individual questions about input. Registration for each workshop has been limited to 15 attendees in order to allow instructors to give participants effective individual attention. Ideally, institutions send one or two staff members to the training who will then, in turn, be able to provide local training to additional staff at their respective institutions. Finally, a detailed workshop script, along with a standard schedule and set of handouts, was developed for in-person training so that other Task Force members and staff at contributing institutions can provide training sessions as needed.

In addition to in-person workshops, online training modules that will be accessible on the Web are being developed. The online modules will allow users to engage in individual, informal training as needed.

Outreach to Non-collegiate Institutions

The Task Force remains committed to making the encoding tool and finding aid repository accessible to as many institutions in Ohio as possible. During 2008, its core base of users (OhioLINK members and academic institutions) was primarily targeted, and in 2009 broader outreach to historical societies, public libraries, museums, and other cultural heritage institutions will be undertaken. This outreach may initially take a more basic form with less emphasis on the EAD FACTORy itself and more discussion of the value of finding aids in general. It will also be emphasized that it is acceptable to produce basic, collection-level finding aids if resources do not allow for more detailed description. In reaching out beyond the OhioLINK community, the Task Force hopes to enlist the assistance of professional organizations such as the Ohio Library Council, which primarily serves public libraries, and the Ohio Association of Historical Societies and Museums. The assistance of early adopters in the public library and historical society/museum communities might also be enlisted to help encourage participation by institutions not currently utilizing the program.

End-User Testing

To date the Task Force has not engaged in end-user testing of the Repository. It will be important to monitor use of this resource and solicit user feedback in order to improve the interface and search functionality from the perspective of researchers. The group has received informal feedback from the OhioLINK Users Service Committee and is currently taking this feedback under consideration for the improvement of the Repository site. One common concern both to end-users and staff at contributing institutions is that digitized content from archival collections be effectively linked to finding aids. The encoding application allows for links to be inserted to digital archival objects in many different areas of the finding aid, including any component-level entry created, but it remains to be seen whether end-users understand the relationship between finding aids and the digital files being accessed. A more basic concern is whether end-users understand what a finding aid is and the extent to which they are able to utilize these descriptive tools effectively.[26] This concern is especially applicable to precollege and undergraduate student users who, having had little or no exposure to archival research, may not be aware of finding aids and their functions.

Tool and Repository Enhancements

As with any application or database, enhancements and additions need to be made on a regular basis. This has become clear as increased numbers of institutions with a wide range of collections and descriptive practices have participated in the project. Enhancement requests are submitted and documented via the Finding Aid Creation Tool itself, and are reviewed on a case-by-case basis by the development team. Meanwhile, institutions already registered for use of the Finding Aid Creation Tool have been encouraged to begin populating the Online Repository, as that resource will only

be useful if a significant number of finding aids from a large number of institutions are included. Several institutions holding "legacy" EAD finding aids have uploaded content into the Repository, and newly created finding aids are being added on a regular basis as well. To date there are 470 finding aids in the Repository.

The Task Force's work since the project's initial release has also included enhancements to the Online Repository. In 2009, XTF 1.9 was succeeded by an upgraded version, 2.1.1. A Task Force member worked with an OhioLINK senior systems developer to implement the upgrade. During the months preceding the anticipated upgrade, the Task Force asked its earliest implementers to send problem reports and requests for enhancements, so that the upgrade, bug fixes, and improvements could be integrated into one process, with the results of all three released at the same time. Implementers' requests centered around issues such as more effective displays of data elements with complex, multipart structures; supplying or correcting missing labels and headings; adding clearer navigation links to the searching and browsing interfaces; and programming an expanded search ability that would find not just subject headings but personal and corporate names, genre and format terms, and other controlled access headings regularly encountered in EAD. Some of these requests were handled in the upgrade of XTF itself; others were locally programmed into the code configured for the Online Repository. The improved Repository was announced to Ohio's EAD community in May 2009, with a specific list of enhancements sent to finding aid contributors.

Finally, the Task Force is monitoring changes to the structure of EAD itself as it has recently been moved to a schema structure. The timing of conversion of finding aids created by the encoding tool to this new structure is still under discussion.

Conclusion

The OhioLINK EAD project has allowed institutions in Ohio to create EAD finding aids without having to invest in extensive training for staff or proprietary software to load locally. The ease of registration, the user account set-up, and the accommodation of multiple levels of finding aid encoding (from very simple, collection-level documents, to detailed, hierarchical descriptions of collection content) offered by the Finding Aid Creation Tool make it highly accessible and appealing to the archival community. Feedback from current users and those exposed to the tool during workshops has been overwhelmingly positive. Perhaps the most promising aspect of the OhioLINK EAD program, however, is the publicly accessible Online Repository, which, if effectively populated with data, will prove to be an indispensable resource for students and other users seeking access to primary sources, manuscripts, and other archival materials. The advantages of being able to conduct a search of multiple institutions' holdings in one location cannot be overstated, as there is currently no easy way to accomplish this.

The concept of a statewide finding aid repository is, of course, not a new idea. The Task Force took much of its inspiration for Ohio's repository from similar programs, such as the Online Archive of California and the Five Colleges Archives and Manuscripts Collection site. As is likely the case with most projects of this nature, the success of the OhioLINK EAD project is due, in large part, to the commitment made by key

OhioLINK institutions, staff members serving on the Task Force, and others around the state who provided testing of and feedback on components of the project. This consortium-driven effort is yet another example of the cooperative environment that has become a hallmark of the library community in the state of Ohio.

ABOUT THE AUTHORS: Cara Gilgenbach is head of Special Collections and Archives at the Kent State University Libraries. She has been a member of the OhioLINK EAD Task Force since 2004 and has served as its chair since 2007.

Amy McCrory is the digital imaging specialist in the Preservation Department of The Ohio State University Libraries. She received her M.L.S. from Clark Atlanta University in 1995. She has served on the OhioLINK EAD Task Force since its inception in 2004.

David Gaj is a Web developer for University Libraries at Kent State University. He received his B.S. in computer science from Clarion University of Pennsylvania in 1988. David did the software development for the Finding Aid Creation Tool.

NOTES

1. "The Ohio Library and Information Network," October 3, 2008, http://www.ohiolink.edu/about/what-is-ol.html (accessed May 15, 2008).
2. "OhioLINK Builds Digital Resource Commons," *Advanced Technology Libraries* 35:8 (August 2006): 6–7.
3. Tom Sanville, "OhioLINK: A U.S. Resource Sharing Facility—Issues and Developments," *Interlending & Document Supply* 35:1 (2007): 31–37.
4. More recently, EAD has been made available as a schema—another way of expressing the same instructions found in the DTD. This article deals exclusively with the DTD version of EAD.
5. Elizabeth Yakel and Jihyun Kim, "Adoption and Diffusion of Encoded Archival Description," *Journal of the American Society for Information Science and Technology* 56:13 (2005): 1427–1437.
6. For information on the EAD Task Force rosters from 2004 to the present, see page 3 of the EAD Starter Packet, accessible online at http://platinum.ohiolink.edu/dms/ead/contentguide/EAD_starter_packet.pdf.
7. The Task Force used the Research Library Group's *Best Practices for the Use of EAD* guidelines as a model for the development of content standards. The RLG guidelines are available at http://www.oclc.org/programs/ourwork/past/ead/default.htm. The OhioLINK EAD content standards are available at http://silver.ohiolink.edu/dms/ead/contentguide.
8. For example, within the academic community OhioLINK serves, the preferred citation field was deemed to be of importance to users. Therefore, it was made a required data element for OhioLINK EAD finding aids even though it is an optional element in the EAD standard.
9. Elizabeth J. Shaw, "Rethinking EAD: Balancing Flexibility and Interoperability," *New Review of Information Networking* 7 (2001): 117–131.
10. A number of software template options were explored, but most involved the end-user needing to know and understand, on some level, EAD XML tagging. The Task Force decided that if actual EAD tagging was involved in the interface, many institutions would not be able to absorb the technical training and support required in the use of such applications.

11. In EAD, collection-level description applies to the entire collection of materials being described, while component-level description focuses on portions or components of the collection, such as series, sub-series, files, and items. Component-level description is optional, but is commonly included.
12. Sonia Yaco, "It's Complicated: Barriers to EAD Implementation," *American Archivist* 71:2 (Fall/Winter 2008): 456–475.
13. Starting in 2005, the Task Force made an effort to provide a presentation at one or more statewide conferences in order to gauge interest in the project, solicit feedback on the needs of the archives and special collections communities in Ohio, and spread the word about the goals and accomplishments of the project. Twice the Task Force members presented at the Academic Library Association of Ohio annual conference, and also provided presentations and a workshop at the Society of Ohio Archivists annual meetings. In addition, Task Force members presented on the project at a number of smaller organizations' quarterly meetings. Interactions at these meetings produced a great deal of important feedback on potential users' needs.
14. Universal codes that designate the language(s) used. See Library of Congress Network Development and MARC Standards Office, "Registration Authority: ISO 69-2," *Standards at the Library of Congress*, July 1, 2008, http://www.loc.gov/standards/iso639-2 (accessed February 19, 2009).
15. EAD tagging includes optional MARC encoding analogs that tie EAD fields to their MARC counterparts. These encoding analogs can then be utilized to more easily generate a MARC record for an archival collection.
16. Additional information about the application's technical structure can be found in the project's Technical Overview document, http://platinum.ohiolink.edu/dms/ead/EAD_Tool_Technical.pdf.
17. Current user levels allowed in the Tool are read-only, read/write own, read/write all, and admin. A special user account (super-admin) is assigned to project leaders who are administering the Tool and can create institution accounts and perform other maintenance and enhancement functions. In this article we have referred to a super-admin account holder as the "EAD Administrator."
18. XSL is typically used to transform an XML document into HTML or other formats suitable for viewing via the Web, PDF Viewer, and so forth.
19. When utilizing this option, encoders are warned that the "suppressed" data will continue to exist in the XML file itself, so use of the finding aid file on a local Web site or other venue will not include display suppression for these fields. However, the data will not be displayed in the OhioLINK Finding Aid Repository itself.
20. Sheila Yeh, Senior Systems Developer at OhioLINK, has completed the XTF programming for the Finding Aid Repository.
21. Archivists' Toolkit™, "List of AT Users," http://www.archiviststoolkit.org/support/ListofATUsers.htm (accessed July 21, 2009).
22. Archon™, "List of Implementors," http://www.archon.org/implementors.php (accessed July 21, 2009).
23. Society of Ohio Archivists, "Merit Awards," http://www.ohiohistory.org/ohiojunction/soa/merit.html (accessed July 13, 2009).
24. See *EAD Application Guidelines for Version 1.0*, Chapter 3, especially Section 3.3.2, http://www.loc.gov/ead/ag/agcre1.html (accessed February 12, 2009).
25. Ibid., Section 3.3.1.
26. C. J. Prom investigates this question in "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 71:2 (Fall/Winter 2004): 234–268.