A Long and Twisted Road: The Journey from EAD to ArchivesSpace Implementation at the University of Minnesota

Abstract:
This article will provide an overview of the implementation of EAD by the department of Archives and Special Collections Department of the University of Minnesota Libraries and the decision to implement ArchivesSpace, which necessitated addressing divergent and legacy practices. Some attention will be given to previous efforts to standardize description and accessibility at the University of Minnesota and how those efforts ultimately failed without a centralized content management system.

Authors:

Kate Dietrick
Archivist
Berman Upper Midwest Jewish Archives, University of Minnesota
320 Elmer Andersen Library
222 - 21st Ave S
Minneapolis, MN 55455
USA
diet0134@umn.edu

Lara Friedman-Shedlov
Description and Access Archivist
Kautz Family YMCA Archives, University of Minnesota
318 Elmer Andersen Library
222 - 21st Ave S
Minneapolis, MN 55455
USA
ldfs@umn.edu (to whom correspondence may be directed)

Caitlin Marineau
Assistant Curator
Children's Literature Research Collections, University of Minnesota
113 Elmer Andersen Library
222 - 21st Ave S
Minneapolis, MN 55455
USA
cmarinea@umn.edu

Introduction

There is perhaps no innovation in the archival profession that has done more to advance the accessibility and discoverability of archival description in the United States than the development of Encoded Archival Description (EAD). First released in 1998 and more broadly adopted following the release of EAD 2002, EAD ushered the archival world into the internet age by providing a standard for archival description in a way that makes it machine readable, parseable, and shareable. After nearly twenty years, the standard is by no means universally embraced,[1] but is widely accepted as the common currency for sharing finding aids.[2]

The Archives and Special Collections (ASC) Department at the University of Minnesota Libraries began planning for EAD in late 2002 and did not fully implement the standard until 2004. Despite being a large and relatively well-resourced institution in close physical proximity to the Minnesota Historical Society, one of the organizations that helped pioneer the standard,[3] the University of Minnesota Libraries was not among the earliest adopters of EAD. Until they were brought together in a new building that opened in 2000, the fifteen collecting units that form ASC were scattered in various facilities around the campus, in some cases miles apart from each other. Each had its own staff, its own policies and procedures, and its own priorities and levels of technological expertise. With the opening of the Elmer L. Andersen Library, most of the collections came together both physically and administratively for the first time, and centralization of various services and processes became not only feasible, but increasingly compelling.

Legacy descriptive practices were quickly identified as an area that needed attention. Previously, the collecting units functioned more or less as separate and independent repositories. Each had developed its own workflow, and conventions, for finding aids as well as its own methods (or lack thereof) for making them accessible to researchers. While a few units had made some small inroads with EAD, none had developed systematic procedures for its implementation. Very few finding aids could be found online, and there were no tools for searching across them, either within individual ASC units or across archival collections at the University of Minnesota.[4]

A wide variation in style, format, and content of description is unfortunately typical among archival repositories in general, and is an issue with which the archival profession continues to struggle. Varied description creates usability challenges for archival researchers around the world, not just at the University of Minnesota. Lack of consistency only adds to a cognitive load that is already significant when working with materials that are typically diverse in nature and origin.[5] However, with ASC collections now located in the same physical location, greater consistency became imperative. Implementation of EAD emerged as a project around which the ASC units could focus efforts to increase standardization, usability, and access to information on the collections.

EAD Implementation at the University of Minnesota

In their article, "Tales from the Shoulders of Giants: Collaborative Implementation of Encoded Archival Description at the University of Minnesota Libraries," Leslie Czechowski and Lara Friedman-Shedlov describe the planning and implementation of EAD at the University of

Minnesota in detail.[6] Based on the findings of a survey of descriptive practices conducted at the outset of the implementation project, the team knew that existing description ranged from detailed, item-level inventories to collection-level abstracts in formats ranging from word processing documents to spreadsheets to databases (in addition, of course, to non-digital paper-based finding aids). The survey found that each collecting unit used its own format for finding aids and its own wording to explain procedures and rules for access and use of materials.   One unit had a practice of embedding contents listings in scope and content notes rather than listing them in a separate section. Others kept biographical information about record creators in separate descriptive tools. Very little of this description was online or structured in a way that would facilitate searching

The EAD implementation process included the creation of best practices, encoding guidelines, and templates that established a baseline consistency for ASC archival finding aids for the first time. The EAD implementation project manager became the first person hired to fill an ASC-wide position, an important step in adjusting for a future that would include more centralized services for the department. As noted by Czechowski and Friedman-Shedlov, although many repositories that were struggling to implement EAD at that time, prioritized finding or creating tools to aid in the encoding process, we found that the encoding proved less of a technical challenge to our staff than finding a tool that would allow users to search across and access the finding aids online.

Software like Archivists Toolkit was still under development at the time, and we did not believe its initial iteration would easily accommodate the level of complexity our descriptive practices reflected.[7] Furthermore, we found that the learning curve for EAD was gentle enough

that staff, including both professionals and student workers alike, could learn to use an XML editor relatively quickly.

With the help of templates and prompts for required content, written guidelines, and a combination of group and one-on-one training sessions, we were able to train a large number of staff across the ASC units to be proficient with EAD within a few months. Over time we also developed processes to facilitate conversion of spreadsheets to EAD box or folder lists, which enabled student staff with little or no training in EAD to contribute to the process. Over the course of the next dozen years, we successfully encoded over 5000 EAD finding aids, with the bulk created in the first few years of legacy finding aid conversion.

Templates and guidelines not only simplified the production of EAD finding aids, but also imposed greater consistency in descriptive practice by enforcing a minimum set of descriptive elements and providing boilerplate text for key information such as access and use conditions. Since EAD focuses on structuring the content of finding aids and does not provide a mechanism for display, we were free to impose a uniform layout and look. Providing a more standardized experience for finding aid users meant a smaller learning curve for users of materials from ASC's multiple collecting units.

With encoding itself proving to be less of a barrier than initially anticipated, a decision was made early on to focus on the acquisition of a discovery interface. The purchase of DLXS, an open source digital library tool developed by the University of Michigan in the early 2000s, allowed researchers to search across all ASC holdings described with finding aids for the first

time..[8] Uniform stylesheets gave the finding aids a consistent appearance and allowed the user to select from several ways to display finding aid information according to preference and need (a structured versus a full text view, for example).

Post-EAD Implementation Issues

After implementing DLXS and going live with the finding aids, we identified many issues with the interface, ranging from display to concerns with search functionality. However, over time, the momentum of the project slowed. Because of other priorities across the Libraries and its IT department, there simply wasn't enough institutional commitment to ongoing enhancement of DLXS. So, from 2007 until our move to ArchivesSpace in 2016, there were very minimal upgrades or fixes to DLXS. A 2009 project to "re-envision" finding aids resulted in a robust set of recommendations and design mock-ups for online search and display of archival description, but they were never implemented because, even then, we anticipated migrating to a new system. We were forced to live with a system that was never perfect in the first place and was becoming increasingly outdated. By 2014, DLXS was no longer maintained or supported. In 2016 our IT department informed us that DLXS was on a server that could die at any moment. We had to find a new system.

An additional frustration was the inability to make changes to finding aids in real-time. Once an EAD file was created or edited, we had to then save the XML file onto a server (another bottleneck, since access was limited to only a few staff members) and then notify the Libraries IT department, which would then re-index the entire folder containing thousands of finding aids. These updates were a relatively low priority for the IT department, which was preoccupied with

administering larger Libraries catalogs and databases. The work was delegated to a single staff member, so there were times where it would take weeks (or even months) to make these changes. In the meantime, staff would have to track which EADs they had updated or added and then follow up months later to see if the new or updated EADs were, in fact, uploaded and available online. This workflow made updates and editing unnecessarily slow and frustrating, and meant that small changes were often put off or ignored for long periods of time. The simple appeal of real-time editing offered by ArchivesSpace was not lost upon us.

Some of the biggest issues we encountered with DLXS stemmed from the workflow management and workarounds we created to address its limitations. One of our largest concerns was that ArchiveGrid and other search engines could not index or harvest data from DLXS, so the finding aids in that system could not be retrieved in a Google search. Our solution to this problem was hosting the XML files on a separate web-accessible server that could be crawled by search engines. From a file management perspective, this proved problematic; our workflow process required generating multiple file copies as the EAD went through the various steps in the process between creation and being added to the web-accessible server space. Staff had to keep track to make sure all copies were identical and up-to-date. Furthermore, because the web-accessible server versions were available online immediately, unlike DLXS versions, some ASC units linked directly to these files from their web pages, rather than to the versions in DLXS. This created a situation where two different versions of a finding aid might be available online. .

Other workarounds operated on a smaller scale. These were ways in which we had to change our work in order to make EADs render the way we wanted in DLXS. One example was

our constant use of <note>. Due to persistent issues with the stylesheets that controlled display, <scopecontent> would not always show up in deeper levels of the description hierarchy (e.g. <c02>, <c03>, etc.) when viewed in DLXS. As a solution, we substituted <note> where we would have normally used <scopecontent>, since <note> was consistently displayed. While this strategy isn't technically incorrect, it is not best practice to use general tags where more specific designations are appropriate. Furthermore, the problem with these smaller workarounds was that, by tweaking the EADs to fit the needs of DLXS, the EADs themselves became somewhat messy and inconsistent, making any sort of migration down the road more difficult.

As noted before, best practices guidelines and finding aid templates were shared with each ASC unit to aid in the creation of uniform EAD documents. As each of our fifteen units had primarily conducted their work in isolation, using templates to guide our work was important. Despite these efforts, as the years went on, the uniformity of the templates began to drift. Once the EAD implementation team that created the EAD best practices documents had completed its work and disbanded, no provision was made for ongoing oversight. While the department established a central archival processing unit in 2010, staffing was limited and its work initially focused primarily on collection storage, space management and addressing processing backlogs in the various collecting units. It did not have the capacity to undertake oversight of descriptive practices or EAD training.. ASC units began subtly tweaking templates or adapting them to address new descriptive challenges that arose. Initial mistakes with templates (for example, the way in which extent was formatted) were not changed across all templates. Without any ongoing individual, group, or system to oversee EAD creation across the department, decisions were made without an awareness of what other ASC units might be doing. As time went on, small

changes were magnified, the differences between individual units' EADs became more and more apparent, and the existing documentation became increasingly outdated and less relevant.

An ongoing issue for the department was training new staff. Not everyone who came to ASC had familiarity or training with EAD and, even if they used templates to create new finding aids, there was no quality control for completed EADs. Thus, incomplete information or errors slipped through. Even for staff who had a strong background in EAD, the tendency to drift over time was nearly unavoidable. Furthermore, the large variety of software tools that people used to create EADs provided room for error..

While some of the issues with our previous workflow were technical in nature, some of them were due to the lack of coordination and cohesion among staff. Different units used different techniques to render EADs, staff strayed from best practices and templates, and there was no central oversight. The history of our department -- with each unit physically separate across campus working in different ways -- created a certain degree of cultural aversion to centralizing work. This may have been reinforced by what Jill Tatem identified in an early study of barriers to EAD implementation as a problem rooted in a professional culture that has not historically enforced "broad or deep habits of collaboration."[9] Without any way to institute or enforce change, ASC units tended to continue to do things "the way they always have." On top of these issues, ASC staff, like archivists in most other repositories, were also struggling with resource limitations, especially high workload and insufficient staffing.[10]

Therefore, as we looked to a new way to create and host EADs, the appeal of ArchivesSpace was immediately evident, being one centralized tool with a structure in place that would, we hoped, steer uniformity of work and output.

ArchivesSpace Implementation

As we began exploring the next phase of archival description, we had significant support and buy-in from our libraries administration. In the years since EAD was implemented, ASC continued to make slow but steady progress towards greater cohesion and centralization as a department. The hiring of the first ASC director in 2005 was a major step. Another was the creation of the central archival processing unit, although it had limited resources and its roles for the first several years extended only to assisting units in processing tasks, rather than creating policy. By 2013, aware that development and maintenance of DLXS would soon cease, the University Librarian committed the Libraries to sustained charter membership in the ArchivesSpace community. Though the University of Minnesota was not among the first to actually implement the program, we were among the 54 charter members who joined in the spring of 2013. In April of that year the department experimented with a demo version, but felt it was not quite robust enough for our purposes, particularly the public user interface that would need to replace DLXS. In 2014, we hired a new head of Archival Processing who came to us with previous experience with ArchivesSpace. This, along with increased staffing for processing, centralized authority to create and enforce cross-unit processing policies and best practices. Additionally, plans underway to implement a major upgrade to the public user interface positioned us to begin work in earnest towards implementing the new system. A long period of

discussion, contract negotiation, and training followed before we began ingesting our finding

aids into ArchivesSpace in the spring of 2016.[11]

Despite the issues that had emerged from the processes and procedures we used, a decade

of working with EAD positioned us well for migration to ArchivesSpace. As a standard ingest

format for ArchivesSpace, EAD finding aids could be imported relatively efficiently using

automated processes. In contrast, repositories whose finding aids were not in a structured data

format like EAD were forced to use a much slower and more labor intensive cutting and pasting

process.[12] Although turnover in staff and lack of ongoing training and enforcement of best

practices caused encoding practice to veer from our original policies, we found that most of the

errors and inconsistencies could be addressed with batch processes using scripts, thanks to the

highly structured nature of the data.

We found that the most common type of error was empty elements that ArchivesSpace

does not allow, including empty date fields and digital object elements without titles. These

issues could be resolved by batch editing the files to replace empty date elements with ones that

contained a generic date (we used "1111-11-11") and adding boilerplate text as the titles of

digital objects. Along similar lines, we found that the contents of certain elements such as

<controlaccess> needed to be formatted a specific and consistent way in order for the EAD files

to successfully import, and many of these could also be corrected using scripts. While there were

some types of errors and issues that ultimately needed to be addressed manually by unit staff, the

amount of hands-on editing was greatly reduced by having finding aids structured in a way that

facilitated these automated processes.[13] While powerful tools like ArchivesSpace may now make

EAD less relevant as a tool to drive the search and display of archival description, EAD clearly remains the key standard for transmission and exchange of archival finding aids. Paromita Biswas and Elizabeth Skene reported in their article on the experience of Western Carolina University's Hunter Library, which had only 35 EAD finding aids with the rest in HTML or other formats, that after a year they had migrated 120 out of 700 finding aids.[14] In contrast to repositories that have transitioned to ArchivesSpace without first implementing EAD, our process was quick and smooth, with 5,600+ finding aids migrated over the course of a few months.

Based on our previous experiences with DLXS, when we approached ArchivesSpace implementation, we knew we needed a better plan to ensure ongoing technical support. We therefore negotiated with LYRASIS to become a hosted institution for ArchivesSpace. By doing so, LYRASIS would provide the Libraries with technical support and server space, and run all ArchivesSpace updates as they rolled out. Any bugs encountered or special requests would be handled directly by LYRASIS, rather than locally. Furthermore, as charter members of ArchivesSpace, we could also provide feedback and help steer updates and enhancements in the years to come. We did not need to rely on the ebb and flow of support coming from internal IT staff, but rather contracted with LYRASIS to provide all the help we needed.. Being hosted by LYRASIS has greatly streamlined the publication of finding aids, as well as troubleshooting and resolving issues.

Another advantage of ArchivesSpace is that it allows us to make updates to finding aids in real time. Whether it is a simple spelling correction or a completely new finding aid,

ArchivesSpace allows staff to update information themselves, click save, and see the changes published immediately, rather than having to wait for someone with enough knowledge of EAD to edit the file and then for the entire system to be re-indexed by local IT. This is a fairly simple benefit, but one that has made some of the most drastic improvements to our workflow and ability to be responsive to patron and donor needs. New accruals can be added to finding aids quickly and efficiently, spelling errors can be corrected promptly, and new details and description can be easily added at any time. For many departments, this has made such an improvement in workflow that it has resulted in the ability to significantly reduce longstanding backlogs of processing work and finding aid updates.

Additional changes included the growing authority and effectiveness of the central archival processing unit. Although Archival Processing had been created long before the implementation of ArchivesSpace, this cohesive and centrally administered platform enabled an increased ability to oversee repositories and control the information within finding aids. ASC units no longer had to deal directly with encoding, which reduced training needs and freed up time to create more finding aids and reduce backlogs of hidden collections. Students and processing staff now require little or no understanding of EAD as a standard in order to create finding aids; as long as they have an understanding of DACS they can instead focus efforts on finding aid content rather than encoding. For the majority of paraprofessional processing staff, training in EAD coding would be superfluous, and they would likely quickly forget any training they received, since they would not be encountering the coding on a regular basis (if ever).

Archival Processing also created templates in ArchivesSpace for each unit's finding aids, which help control descriptive practices to a greater degree than was possible while using encoding tools. With ArchivesSpace, since all description is happening in one central system, there is less danger of a new staff person remaining unaware of existing templates. Even if staff alter the templates, the software constrains changes within specific limits, reducing the opportunities for widely differing practices for finding aids across ASC, or even within units.

One of the biggest benefits with the introduction of ArchivesSpace is the ability for the finding aids to be indexed by search engines such as Google. Previously, search engines did not index files in DLXS, so we were required to take additional technical steps to make our finding aids appear in internet searches. The ability for the materials to be immediately available to search engine crawlers increases findability of the collections, and also reduces staff workload.

Finally, the department has been able to centralize other aspects of workflow within ArchivesSpace, to a degree that we had not initially anticipated. ASC units have started documenting new accessions, donor information, and storage locations within ArchivesSpace. These are functions that were previously managed, like description, with widely varying tools and procedures. The ability to use one program to handle nearly all of the documentation for a collection has increased uniformity in overall handling of materials across all of ASC. Though the ASC units still maintain paper files and spreadsheets to document certain aspects of their collections, the introduction of ArchivesSpace has allowed a significant reduction in the amount of information that has to be documented in these ways. In addition to streamlining current

operations, information is now more likely to transfer smoothly between staff members throughout the building or over time when staff leave their positions.

No system is perfect, so there were of course issues with ArchivesSpace itself, as we found during our implementation process. We chose to go live with the public user interface (PUI) in September of 2016 despite knowing that it had a number of serious usability issues, but that a major overhaul was underway. In the time that we waited for these PUI updates, we dealt with annoyances that ranged from the minor, like certain formatting not being allowed in notes fields, to the major, like the inability to view an entire box/folder list. While in most ways, creating finding aids is now easier and more accurate than our previous methods, ArchivesSpace lacks certain basic editing features like spell checking and search/replace. We have also found that it remains more efficient to create certain types of folder lists outside of ArchivesSpace using spreadsheets and import them as EAD files for further editing. Although ArchivesSpace provides the ability to create "rapid data entry" templates that can be used for entering multiple components, this feature has proven to be prone to glitches. Even when the rapid data entry feature is working as intended, some staff find that creating long lists of files remains more efficient with spreadsheet software that includes features like the ability to quickly autofill columns.

Despite issues with ArchivesSpace, we knew as charter members we could potentially influence updates and changes. The updated public user interface released in August 2017 has alleviated many of these issues, and future updates and usability testing should help continue to refine the interface. A downside with a system like ArchivesSpace, where we are hosted by

another company, is that we do not have direct control over the display. However, as our experience with DLXS has shown, direct control can be a problem if you do not have the resources to regularly edit and update.

Conclusion

As with many repositories, the path of the University of Minnesota Archives and Special Collections to standardize archival finding aids for researchers has been a long and twisted one. Though our previous adoption of EAD left us well-poised to transition fairly smoothly and efficiently to an archival management software like ArchivesSpace, our particular approach to EAD created issues that ultimately made the transition to a new system necessary. In the year since implementing ArchivesSpace, we have been able to address many of the issues of consistency and coordination, facilitated greatly by an increased ability of Archival Processing to oversee template creation and robust description. In addition, the ability to use the same software for accessions and location management supports our move towards greater centralization of key functions and processes in ASC, as well as better coordination of and attention to collection management.

Lessons learned from our experience with the transition process, as well as our previous efforts with DLXS, have served us well. They have pointed the way towards what the department needs to focus on in order to maximize usability and reliability for our users and staff. This includes the need for ongoing support and oversight of software and finding aid creation, as well as the need for enforcing the consistent use of templates for creating resource records. Finally, these lessons led us to a new understanding of the role of standards for archival

description. Although we remain committed to EAD as an output for sharing records across

platforms, the ability of staff and students to create finding aids without needing experience in

EAD means that archivists can focus increasingly on content and user access, rather than

technical needs. Description and access have always been a foundation of the archival profession

and a primary focus of where we channel resources. As we continue to refine our workflows

with more powerful tools like ArchivesSpace, we now have better mechanisms for realizing that

goal.

Notes

[1]Karen F. Gracy and Frank Lambert, "Who's Ready to Surf the Next Wave? A Study of Perceived Challenges to Implementing New and Revised Standards for Archival Description," *American Archivist* 77, no1 (Spring/Summer 2014): 96-132. The survey conducted by Gracy & Lambert found that of 324 respondents, 52% said their repository was "using EAD."

[2] Very little in the way of statistics or hard data on current adoption levels is available, but inferences made from the amount and type of discussion in the professional literature and conferences suggest that EAD is considered a mature and broadly accepted standard. See also Society of American Archivists. Technical Subcommittee for Encoded Archival Description, "Preface." *Encoded Archival Description Tag Library, Version EAD3* http:/loc.gov/ead/EAD3taglib/index.html#d0e115; Jenn Riley and Kelcy Shepherd, "A Brave New World: Archivists and Shareable Descriptive Metadata," *American Archivist* 72, no. 1 (Spring/Summer 2009): 91-112.

[3] The University of Minnesota is only a few miles down the road from the Minnesota Historical Society, which under the leadership of Michael Fox, implemented EAD in the late 1990s.

[4] Leslie Czechowski and Lara Friedman-Shedlov, "Tales From the Shoulders of Giants: Collaborative Implementation of Encoded Archival Description at the University of Minnesota Libraries," *Journal of Archival Organization* 5, no. 3 (2007): 9-29.

[5] In 2003, Elizabeth Yakel articulated this problem when she pointed out that, "The very act of archival representation, designed to order and provide access to collections through finding aids, can also create barriers to use. Researchers must know the schemas and codes and understand the underlying systems of privileging, classifying, and selecting that comprise both arrangement and description." "Archival Representation," *Archival Science* 3, no. 1 (2003): 2.

[6] Czechowski and Friedman-Shedlov, "Tales from the Shoulders of Giants."

[7] "Archivists' Toolkit Overview," accessed September 8, 2017, http://www.archiviststoolkit.org/overview/phase2.html.

[8] See http://www.dlxs.org/about/aboutdlxs.html

[9] Jill Tatem, "EAD: Obstacles to Implementation, Opportunities for Understanding," *Archival Issues* 23, no. 2 (1998): 155-169. p.160.

[10] See for example, Sonia Yaco, "It's Complicated: Barriers to EAD Implementation," *The American Archivist* 71, no. 2 (Fall/Winter 2008): 456-475.

[11] For more information on the implementation process we undertook, see Lisa Calahan and Kate Dietrick, "Setting the Stage and Keeping Sane: Implementing ArchivesSpace at the University of Minnesota," *Journal of Archival Organization,* 13, no. 3-4 (2018).

---

[12] See the implementation process described by Paromita Biswas and Elizabeth Skene in "From Silos to (Archives)Space: Moving Legacy Finding Aids Online as a multi-Department Library Collaboration," *The Reading Room* 1, no. 2 (2016): 65-84.

[13] More detail on our migration challenges and solutions in Calahan and Dietrick, "Setting the Stage," 2018.

[14] Biswas and Skene, "From Silos to (Archives)Space," 2016, p. 77.