EAD3 Implementation Survey Results and Discussion
Wim van Dongen and Katherine M. Wisser

Introduction

Since the release of the first official version of the EAD standard in 1998, institutions have struggled with its implementation. Even long after the release of the first revision of the standard in 2002, surveys have exposed the challenges faced by institutions. These challenges primarily focus on the complexity of the technical implementation, especially in smaller institutions with few resources.[1] Therefore, the Technical Subcommittee on Encoded Archival Standards (TS-EAS), thought it necessary to monitor the implementation of the second major revision of the EAD standard – EAD3 v1.0.0 – closely and to accompany the drafting of the first minor revision of this new EAD version, which led to EAD3 v1.1.0 being released in June 2018, with an online survey.

This report provides the result of this EAD3 Implementation Survey as distributed by the TS-EAS between June 2017 and September 2018. Based on these results, follow-up activities are being planned to address some of the major concerns raised by respondents as well as to support and facilitate an easier implementation of EAD3.

Survey design, distribution, and general response rates

The survey was designed for three audiences: institutions, companies or software developers, and individuals. The answer to the first question guided the respondent to the set of questions related to the audience they represented. The questions were designed to assess the respondent's current usage of EAD, the intention of implementing EAD3, and in the case that EAD3 is not being considered, the reasons why. The Technical Subcommittee hopes to be able to address any misconceptions caused by ambiguity in the survey questions and to help direct next steps to facilitate implementation for those experiencing obstacles (such as the identification of needed tools).

The survey was first published in June 2017 and was advertised via several communication channels. Within one month 98 responses were collected. A web-based survey can easily be abandoned at any time, leaving an incomplete result; of the initial 98 responses, only 46 were complete and usable. A preliminary overview of the results of those 46 responses was considered by the Technical Subcommittee, and its members determined that a more complete picture of implementation was needed. In July 2018, the survey was disseminated again with a closing date of September 2018. Combined with earlier responses, the total number of responses was 203 over a total period of 16 months. This results in two periods of activity (see Figure 1). Due to the above mentioned functionality of a web-based survey, only 88 of the 203 responses were complete.
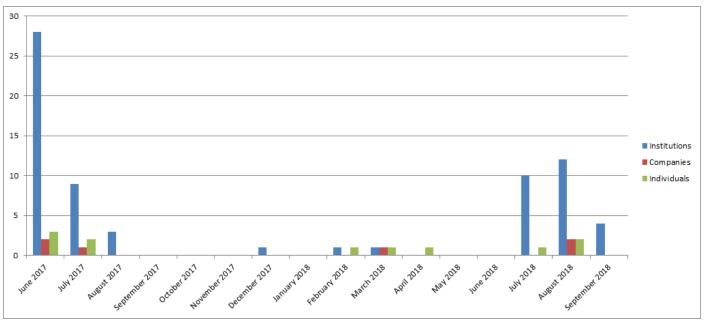

Figure 1: Periods of activity for the survey

---

[1] See: Sonia Yaco, "It's Complicated: Barriers to EAD Implementation" *American Archivist*, 71:2 (Fall-Winter 2008), pp. 456-475, Michele Combs, et al., "Over, Under, Around, and Through: Getting Around Barriers to EAD Implementation" Report produced by OCLC Research in support of the RLG Partnership. Published online at: http://www.oclc.org/research/publications/library/2010/2010-04.pdf, and 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,* 77:1 (Spring/Summer 2014), pp. 96-132.

Nine institutions and companies completed the survey twice. For two cases it is clear that the newer responses, provided shortly after the first submissions, were made to add information. Their first submissions were eliminated. Other duplicates had submission dates that were separated by a year; most, but not all, showed the same information, and some newer answers contained interesting extra information; therefore, they were not deleted. This recalibration resulted in a total of 86 complete and usable responses.

The survey first asked respondents to identify themselves as an institution, company, or individual. Ultimately, 69 institutions, 6 companies, and 11 individuals completed the survey. Of the institutional responses, 46 came from the United States and 23 from Europe (including Austria, Bulgaria, Croatia, France, Germany, Hungary, the Netherlands, Spain, Sweden and the United Kingdom). Of the six company responses, five came from Europe and one came from the Middle East. The survey did not ascertain the geographic location of the individuals who completed the survey under that strand. While the companies and individuals provided some interesting insights that will be discussed later, the most significant findings from an implementation standpoint were those of the institutions. This report will provide a discussion of the specific survey results for each category surveyed and will suggest some steps forward to address any concerns raised in responses.

Specific Survey results

Institution responses

The questions for institutions were divided in two parts: first, there was a series of questions that addressed their current use of EAD; second, there were specific questions about EAD3 and plans for implementation. The first institution-specific question asked whether or not the institution was using EAD and if so, which version. The responses were overwhelmingly skewed to EAD 2002, accounting for over 75% of survey respondents (see Table 1).

Table 1: Use of EAD at Responding Institutions (n=69)

Answer choices	Number of respondents	Percentage
Not using EAD	3	4.35%
Using EAD version 1	0	0.00%
Using EAD 2002	53	76.81%
Using EAD3	8	11.59%
Other	5	7.25%
Total	69	100.00%

Those selecting "other" were asked to specify what that meant. Two of those responses indicated EAD (DDB), which was described as a subset of EAD 2002; another indicated apeEAD which is also a subset of EAD 2002; finally, an institution indicated they were encoding finding aids using a template provided through the Texas Archival Resources Online (TARO) project, which they believed is based on EAD 2002. Given that those four responses could reasonably be considered EAD 2002 implementations, that brings the total percentage of institutions using EAD 2002 to over 80% (82.60%). The last "other" response was ISAD(G) but no further explanation as to that standard was provided.

The next question focused on software tools used in repositories for the creation of EAD files. The options did not ask about specific commercial software tools, but rather asked about the origin of the software tools or the integration into larger systems. Table 2 provides an outline of the responses.

Table 2: Software tools used at institutions for creating EAD files
(n=63, 91.3% of total institution respondents)

Answer choices	Number of respondents	Percentage
Specific software tools developed for and by the institution	8	12.70%
Specific software tools developed for the institution by others	5	7.94%
Exports out of a standard collection management system	20	31.75%
Commercially available XML editor not specific to the institution	20	31.75%
Other	10	15.87%
Total	63	100.00%

These results are well distributed across the various options. This question presented an erroneous assumption that institutions would have a single strategy. This issue was corrected in a couple of ways by survey respondents. The larger number of "other" responses was one way that respondents identified the reality of hybrid strategies for EAD creation. The other way that respondents corrected the problem inherent in question 10 was to select a category but then answer each of the follow up questions, demonstrating the use of a variety of strategies.

When asked about software tools that were developed for the institution (question 11), responses included an XMetal editor plugin, AXAEM, .net SQL database, PL/SQL procedures in an Oracle database, and XML exports out of MySQL/Oracle databases. Many answers in this category also identified a number of tools developed in-house, including software specifically built to support all levels of the archival description, or a combination of strategies involving commercially available XML editors. Also scripts to convert Excel files into EAD were mentioned.

The next two questions asked for the name of collection management systems supporting EAD at the institution and the name of commercially available XML editors in use. For collection management systems (question 12), respondents identified 18 different software packages, including: ActaPro, ArchivesSpace, Archivists Toolkit, Archon, Arcinsys, Arkis, AXAEM, FAUST, FileMakerPro, HEUSS, ISDA, Memorix Archives, Islandora, Ex Libris DigiTool, scopeArchiv, Tessela SDB, and XTF. ArchivesSpace accounted for nearly 50% of the responses (46.88%). Archivists Toolkit accounted for almost 10% (9.38%) and ActaPro and Archon just over 5% (6.25%). All other programs were mentioned by only one institution. When asked about commercially available XML editors used at institutions (question 13), eight different XML editors were identified, including: EAD Report Card, MarcEdit, Notepad ++, NoteTab Pro, oXygen, Sublime, XMAX (JustSystem), and Xmetal. Of these answers, oXygen was the most prevalent (78.38%). Notepad ++ and NoteTab Pro were used by 10.81% of the respondents each and Xmetal by 8.10% of the respondents.

Taken as a whole, despite the ambiguity of the initial question in this group, it is clear that repositories are using a wide variety of strategies to create and manage their EAD files. These strategies range from highly tailored systems that generate EAD XML files specifically, whether commercially available or homegrown, generic XML-specific editors, to generic database and spreadsheet tools with scripts developed to convert data into an XML format. While some solutions appear to be popular in institutional implementations, there is not a single or even a dominant strategy.

Three additional questions address EAD practices in general for institutions. The first two focused on best practices and workflow documentation. Question 14 asked if the institution used best practice or workflow documents for the creation of EAD files (see Table 3).

Table 3: Use of best practices or workflow documents (n=69)

Answer choices	Number of respondents	Percentage
Yes	41	59.42%
No	19	27.54%
Not applicable	9	13.04%
Total	69	100.00%

Question 15 asked that if the institution did use best practice or workflow documents, were they available online. Of the 24 that answered this question, 22 answered affirmative. Five institutions indicated that their documentation was not publicly available (22.73%) and 17 indicated that their documentation was publicly available (77.27%). Two institutions indicated "not applicable," although it's not clear how that category was understood as distinct from the "no" category. The last question in this section of the survey asked about online presentation of the institution's content based on EAD files. Of the 48 respondents, 46 indicated that their content was publicly available (95.83%) and only 2 indicated that it was not available (4.17%).

Taken as a whole, the general implementation of EAD across repositories shows that institutions are still primarily based in EAD 2002, although there appears to be some adoption of EAD3. Of the institutions responding, there was a complete shift from EAD version 1.0, however. Strategies for creating and/or managing EAD files appear to be dispersed across different software platforms, collection management systems, XML editors, and other non-XML-specific tools. Institutions trend toward adopting hybrid strategies, indicating that there is no consensus on the best or most efficient single strategy; this should not be surprising given all the variables that go into the development of implementation strategies. Best practices or workflow documentation does appear to be valued by institutions and is often made publicly available. A follow-up question exploring why an institution would not make their documentation publicly available may be a useful avenue to explore. However, of the 66 institutions that said they were using EAD, 46 indicated that they provided an online presentation of the content based on EAD files. Twenty of those institutions did not respond to this question or may have responded no. An additional three respondents indicated that they did not use EAD, and it is not clear whether or not those institutions were able to answer additional questions in this part of the survey. Because of this ambiguity, some results provided limited insight.

The second section of the institution survey focused specifically on the adoption of EAD3. Question 17 focused on whether or not the institution was planning on using EAD (see Table 4).

Table 4: Plans to use EAD3 (n=69)

Answer choices	Number of respondents	Percentage
Yes, already working on it	10	14.49%
Yes, but not right now	43	63.32%
No	16	23.19%
Total	69	100.00%

The information presented in Table 4 is an overview of all responses over the whole 16-month survey period. A comparison between the two active periods in that 16-month span provides more detail. Table 5 outlines the responses to this question from June - August 2017 (the first period of activity).

Table 5: Plans to use EAD3
First active period for survey: June - August 2017
(n=40)

Answer choices	Number of respondents	Percentage
Yes, already working on it	6	15.00%
Yes, but not right now	25	62.50%
No	9	22.50%
Total	40	100.00%

Table 6 outlines the responses to this question from the second survey push period from July - September 2018 (the second period of activity).

Table 6: Plans to use EAD3
Second active period for survey: July - September 2018
(n=26)

Answer choices	Number of respondents	Percentage
Yes, already working on it	3	11.54%
Yes, but not right now	16	61.54%
No	7	26.92%
Total	26	100.00%

A comparison of these two periods shows only a slight decrease in the interest in implementing EAD3 when looking at the 'Yes' answers: 77.50% in the first survey period compared to 73.08% in the second survey period. Note that an additional three responses were completed sometime between these two periods of activity.

Overall, though, almost 15% of the institutions responding indicated that they were not planning on implementing EAD3. In order to provide the Technical Subcommittee with information about community reluctance or understanding of the standard, a follow up question was asked that focused on those institutions that answered 'No' to the previous question. Institutions were asked to explain the reasons for not moving to EAD3. These free-text answers can be grouped into three major concerns: the cost of change versus the perceived benefit, external system or aggregation dependence, and on-going developments in the archival standards landscape, which creates a sense of instability. Below is a fuller discussion of each of these obstacles.

The challenge to migrating to EAD3 was the most prevalent obstacle indicated in this section of the survey. One respondent noted, "fails the cost-benefit test: too much work to convert [...] our finding aids (not to mention all the supporting scripts and delivery mechanisms that would also have to be adjusted), and the new features are not things we see a need for." The adjunct activities involved in EAD programs came up in other responses: "several inconvenients [sic]: necessity of training again catalogers who are not EAD experts, names of element and attributes renamed with no necessity ..." and "I'd have to rewrite a lot with no real benefit added. It's easier and more effective just to hack EAD 2002 a bit ..." Other responses indicated that they did not have the resources to support an effort at this time: "To be honest, I'm not sure, I'm a lone arranger [...]." Directly tied to the challenges in implementing EAD3 are the perceived lack of benefits in doing so. For example, one respondent stated, "new features are not things we see a need for" or "no need for new feature or functionality added in EAD3." Perhaps most critical is the comment that

4

"some new elements and attributes which are introduced in the schema seem quite unnecessary: they lead to pointless complexity for professionals and are likely to result in creating an unwanted wide gap in practices and renditions of EAD."

Some respondents indicated their program dependence on external factors, whether their EAD creation and/or management system or the aggregation they participate in. Our 'lone arranger' identified above noted "if TARO switches to EAD3, then I imagine, yes, I will be too, but it's going to be driven by TARO, not by me individually at my institution." Another respondent stated "In order to use this, Memorix Archives needs to be adjusted." These responses suggest that external systems or participation in aggregations is a primary factor in the use of the standard. What is not clear is whether this dependence can be seen as a negative impact of the move away from building internal expertise and implementation strategies.

A few respondents indicated that they were waiting for other initiatives in the standards landscape before they migrate any of their descriptive work. One respondent simply stated, "ICA is developing an archival conceptual model and an ontology. This will have consequences on SAA encoding standards, implying that a change at this time would be premature." Another respondent also mentioned this on-going work: "We prefer keeping focused on future ways of exchanging data (e.g., RiC)." The international archival community is eager to see how the work of the Expert Group on Archival Description may have an impact on our communication standards (such as EAD), although it is not clear whether it will directly impact the structure standards or rather the systems that use the data in retrieval and display functionalities.

Those institutions that indicated that they have intentions to implement the standard but are not engaging with it at the time were asked to describe those plans. Many of those responses discussed tentative plans, at best, but again mentioned external management system adoption as a part of that process, while others mentioned the need to migrate to external management systems such as ArchivesSpace as a mechanism for implementing EAD3. Some respondents discussed project prioritization as a factor. Despite the obstacles identified, most responses were positive in tone: "We will move to EAD3 but do not have a plan yet" and: "possibly as soon as next year with other related finding aid development."

Those institutions that indicated that they were already using EAD3 were asked to report on that implementation (see Table 7).

Table 7: Types of implementation of EAD3
(n=12)

Answer choices	Number of respondents	Percentage
Converting existing EAD files to EAD3	3	25.00%
Implementing EAD and using EAD3 to do so	4	33.33%
Collection management system which supports EAD3	4	33.33%
Other*	1	8.33%
Total	12	100.00%

*The other category indicated a completed migration to the new EAD3 standard in 2016.

In order to better understand the institution's implementation processes and uncover strategies already deployed, respondents were asked to provide additional information on conversion software in use for converting existing EAD files to EAD3. Respondents identified local solutions, external management systems such as ArchivesSpace, and tools provided through the EAD website (referred to as the Catapano/Rush stylesheet). When asked about resources for best practices or workflow, local documentation and EADiva were mentioned. Collection management systems were identified as well, including ArchivesSpace, AXAEM, and Omega Software Centrix Pismohrana.

Finally, institutions were asked if there was anything else they would like to share regarding the implementation of EAD3. In the 21 responses to this question, the tone was significantly more positive and optimistic than those in previous sections of the survey. Comments such as "we would love [to]start asap, but..." and "while we wish to take advantage of EAD3's features in the future, we are currently waiting..." and "looking forward to making it happen when we're able." There were also some concrete suggestions provided, including increasing the availability of examples and tools. One respondent suggested "it might be useful to have more persuasive arguments in favour of EAD3. We are not sure that it will materially help with the functionality we provide at the moment."

Company responses

The questions for companies were aimed at obtaining information from them on their plans to start supporting EAD3 in the collection management systems they provide. The first thing to ascertain was whether or not the software provided by the company provides or supports EAD3 already or is planning to do so (see Table 8).

Table 8: Intentions of software to support EAD3
(n=6)

Answer choices	Number of respondents	Percentage
Yes, already working on it	3	50.00%
Not yet, but we are planning to start on it	3	50.00%
No	0	0.00%
Total	6	100.00%

Despite the fact that all company responses indicated that they were either already engaged in supporting EAD3 or planning on it for the future, one respondent noted that that standards seemed "difficult to understand and implement." For those that have already adopted it, one respondent noted that their software is "a recently developed program (2016) so it was decided to use EAD3 for the advantages of using it [...]" Another company noted that they "decided to use EAD3 for its greater ability to represent archival documents on the Web and its relationships with other documents or archives through URIs."

Individual responses

A third section of the survey addressed individuals. The first question asked about their interest in EAD. Reponses include professional involvement in EAD of individual software developers, curators/editors and educators, and general interest. Respondents then were asked if they have worked with EAD 2002. Of the 11 respondents, just over half (54.55%) responded yes. When asked what purpose they had in working with EAD 2002, responses included: web archiving, migration projects, and consultations. A variety of tools were identified as being used in the creation of those EAD files.

When asked about plans for working with EAD3, almost 75% of the individual respondents answered yes. Reasons included being able to respond to need in the community and one noted: "supporting only old standards is foolish." For those respondents that indicated that they were not planning on moving to EAD3, responses indicated a reliance on ArchivesSpace as the primary reason for not moving to EAD3. One respondent said that they were undecided.

Lessons Learned from the Survey and Next Steps

There certainly is an interest in the use of EAD3; however, there are obstacles. From the point of view of institutions these include a lack of resources coupled with a lack of understanding of the benefits of the new version, especially compared to the efforts of moving from EAD 2002 to EAD3. These obstacles can be overcome by cooperation within consortia or aggregators that collect and publish EAD finding aids from multiple institutions and by cooperation with suppliers of collection management systems. The confidence in ArchivesSpace is illustrative of that strategy and that is of course further fueled by the fact that ArchivesSpace already supports export of valid EAD3 for almost a year.

Feedback from the survey demonstrates that there is a need for more dialog about the benefits of EAD3. Both the institutional and individual respondents revealed a perception gap in the advantages to the new coding capabilities invested in the new version. A possible start to address this could include descriptions of use cases for EAD3 extensions, examples of EAD3 files that illustrate data tagging in more meaningful ways, roadmaps from EAD3 towards Linked Open Data models and the emerging Records in Context (RiC) conceptual model. Other strategies may include Webinar formats that explain the changes in detail and offer time for questions from the community to clarify any aspects of the version that may appear unclear.

Finally, there were some comments in the survey that call for additional information gathering. And, of course, this survey represents a snapshot of implementation in time: things can progress very fast in this arena. The extended length of time that this survey was open provides evidence of the speed of change. For example, in July 2017, one institution had noted an intention to implement the standard in their response, but by August 2018, it answered the same question with "already working on it." Therefore, it will be important for the Technical Subcommittee to continue information gathering and surveying efforts in order to be responsive to the needs of the encoding communities. It is equally important for institutions to be willing to share their migration stories, the tools developed, and the best practices or workflow documentation to the larger community.

References

Combs, Michele, Mark A. Matienzo, Merrilee Proffitt, and Lisa Spiro. 2010. Over, Under, Around, and Through: Getting Around Barriers to EAD Implementation. Report produced by OCLC Research in support of the RLG Partnership. Published online: http://www.oclc.org/research/publications/library/2010/2010-04.pdf.

Gracy, Karen F. 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*, 77:1 (Spring/Summer 2014), pp. 96-132.

Yaco, Sonia. "It's Complicated: Barriers to EAD Implementation" *American Archivist*, 71:2 (Fall-Winter 2008), pp. 456-475.