

Final ACERA Meeting Minutes, 11/5/09

**ADVISORY COMMITTEE ON THE ELECTRONIC RECORDS ARCHIVES
MEETING NO. 9
NATIONAL ARCHIVES BUILDING
MINUTES
DAY 2 OF 2
NOVEMBER 5, 2009**

In accordance with the provisions of Public Law 92-463, the meeting was open to the public from 9:00 a.m. to 5:00 p.m.

COMMITTEE MEMBERS

<u>Name</u>	<u>Organization</u>
Laura E. Campbell	Library of Congress
David Carmichael	Georgia Archives
Sharon Dawes	Center for Technology in Government
Dr. Richard Fennell	Administrative Office of the U.S. Courts
Dr. Christopher Greer	National Science Foundation
Jerry Handfield	Washington State Archives
Robert Horton	Minnesota Historical Society
Dr. Robert E. Kahn	Corp. for National Research Initiatives
Andy Maltz	Academy of Motion Picture Arts and Sciences
Dr. Robert Martin	Retired Endowed Chair Texas Woman's University
Martha Morphy	Assistant Archivist for Information Services
James Neighbors	U.S. Air Force
John T. Phillips	Information Technology Decisions
Daniel V. Pitti	University of Virginia
Jonathan M. Redgrave	Nixon Peabody LLP
Dr. Dan Reed – Not Present	Microsoft
David Rencher	Federation of Genealogical Societies
Adrienne Thomas	Acting Archivist of the United States
Dr. Kelly Woestman	Pittsburgh State University

Final ACERA Meeting Minutes, 11/5/09

1. Comments from the Chair

Picked up discussion of marketing and communications working subgroup and determined that issue would be discussed by the committee as a whole initially.

2. Recap of Presentation Regarding Issues of Architecture -- John Phillips

John reviewed his presentation that had previously been given prior to the Chair joining the committee.

A guest, Mike Taft, was introduced.

John briefly reviewed history of database, creation of network and records management applications. Explained how to gain electronic records management control of electronic records and pointed out the problems of increased numbers of users slowing responsiveness of system. Moving large volumes of records across these networks creates a huge repository that needs a large system backup which takes progressively longer to backup. Solution suggested was to leave assets in their original creation applications, but to prevent deletion and to classify the information for purposes of searching. Control is achieved with a permission metadata server.

You need integration, connectors for different kinds of application; creation application and storage application. Gave an IBM example and Microsoft sharepoint.

Chair asked how to query the server for access to data, how to scale the architecture to handle volume. Chair also asked whether you could have an ERA-type architecture that points data back to point of origination. John explained that idea is to leave the records where they originated and control data management from there, rather than move the data to a centralized location with an integration path or toolkit.

Ideally, once an integration path is developed for a certain environment, you don't have to recreate that, and if built on the open-systems architecture of ERA it might be more inexpensive to migrate it. Challenge is integrating with whatever is currently out there. Smaller organizations might benefit from a NARA-initiated development aspect of ERA.

Reviewed challenges of the CLOUD which changes accessioning, and now lawyers do not want this data on the CLOUD because there's no way to authenticate a record.

Report on European Projects

Chair asked Daniel Pitti to report on the international conference he recently attended. Daniel commented on two projects funded by the European Commission. Both projects are exclusively geared toward access, not curatorial preservation, one is European and the

Final ACERA Meeting Minutes, 11/5/09

other is the Apenet, the archive portal-Europe. The cultural underlying theme is to create a common European identity. Some popular critique of Europe likens it to Google.

3. Briefing on Conceptual Framework for Digital Preservation

Archivists David Kepley and Michael Carlson started with previously distributed NARA Conceptual Preservation Framework. Stated that purpose of ERA is ingest and store data sets. Agencies need to be clear on requirements before they can instruct Lockheed Martin in building ERA, hence construction of NARA Prototype. Once consensus was reached in NARA, we needed to do outreach to other entities like ACERA.

Requirement of OIS compliance gave framework for archival information package, submission information package -- accessions from creating agencies -- and dissemination information packages, which are submitted to the researchers.

Archival information package becomes the intellectual entity to be managed through ERA. We are managing more than digital objects, i.e., individual pages or pieces of data, but must managing intellectual entities which may be more than one digital object, the document as a whole, for example.

For original assets, focus is to extract metadata from the submission information package, as it comes from the originating entities, and to construct appropriate archival information packages. Once metadata is captured, it is permanently associated with the digital files that constitute the AIP.

For derived assets want to make sure that an AIP is created for the new asset that will be a combination of AIP information that inherits from the original. Premise is the metadata model used to capture.

Authenticity challenges being met with common-sense approach. NARA must be able to document their authentication processes and practices, must capture and maintain the AIP metadata. Four tests of authenticity: no corruption, be able to document your processes, maintain the IP, and maintain significant properties.

Corollary to this is access the goal is that NARA will make digital files reasonably accessible to researchers and the public. Question was posed as to whether partnering with industry would be a more efficient way to progress towards architecture solutions.

Transforming records is the most expensive strategy, but inevitably there will be materials in esoteric formats that researchers will not have access to, and will need to be transformed.

Final ACERA Meeting Minutes, 11/5/09

Michael Carlson discussed transformation, key question asked when transforming is, are members of the research community who are likely to want to use the records reasonably able to render the records and use it to meet their research needs. Two drivers for this prong are preservation and reference. The hope is that there are not obsolete formats in their holdings and that these are able to be transformed. A scale has been developed for rating the obsolescence to determine difficulty in transformation. The question is can NARA afford to, or should they bother, to transform those objects closest to being obsolete. It was suggested that private industry may be able to research the number of all possible formats that exist.

NARA contemplates building a staff of IT specialists that has task of working with archivists to identify those formats that will be most challenging, and to assess the needs of the research community for certain formats. This would be a collaborative process, determination would be recorded into a reference-driven transformation plan, the ERA system would execute the transformation on specific groups of records. Transformation for preservation may not be the same as the strategies for research and access. In no case should the original be discarded.

Open Office and some of the new programs may ameliorate the transformation difficulties. Question was asked about what the demographic of users, rather than researchers, may look like in the future. The users and researchers are considered to be the public. Users have evolved from purely academia, to now genealogists and possibly in the future photographers and filmmakers, and if metadata is not accessible or useful to broader audience, concern is that there will be many complaints of usability.

Michael showed a sample transformation plan that would take into consideration all of the tools used in the transformation, and effectiveness of the tools. Sustainable format is so important that NARA has promulgated guidance regarding sustainable formats to the Federal community. This will ensure that this process does not have to continually be undertaken. The next challenge will be how transformation will affect authenticity.

Finally, archivists must make significant property determinations at the point of transformation or at the aggregate level, and trying to automate this process as much as possible. Classes of significant properties are appearance characteristics determination, the structure of the records, for example, are all attachments to email included. Idea is that these standards will meet legal requirements of authenticity.

Metadata regarding transformed version would always be associated with the original version and the original catalog entry.

Final ACERA Meeting Minutes, 11/5/09

Jim, who represents the classified environment of the government, pointed out that he would like to collaboratively identify classified materials, rather than current process, which is serially.

Kelly asked if there has been collaboration in developing the system across agencies. The jury is still out on the question of whether NARA should be in the business of managing electronic records.

Question was posed at to whether there would be transformation of executable programs. The problem is getting all of the agencies to give NARA their materials on sustainable formats, otherwise the cost of transformation is passed along.

Also queried whether there was any revenue model for ERA itself from the point of view that revenue could come over to ERA from what is generated from storage of paper records. NARA representative explained that they are exploring the possibility of what makes sense from an Electronic Federal Records Center standpoint. Instead of charging for paper storage, it is being explored to charge for storage of electronic records in something like ERA. Charging for processing these materials may also serve to weed out what is not historically significant data.

John commented that there would have to be a model for a revenue-generating system and asked Andy what he had heard from the transformation topics presented here that were the same for the motion picture industry. Response was that every time transformation is done, something is lost.

Department of Defense has driven many of these standards because of their unique position of dealing with so many private and government vendors.

Presentation was concluded by asking for ongoing feedback from the committee.

<Break 10:30>

<Return from Break 10:45>

4. Roundtable Discussion of Marketing ERA

Chair reiterated committee would not break into working groups, but would be having a roundtable discussion to develop suggestions for NARA's communication strategy about ERA. Need to identify and segment the relevant audiences. Archival community is one segment of the audience further subdivides into State Archives and other subgroups, Congress, the White House and may be slightly different and nuanced messages to each audience, but those messages can not be internally inconsistent. Suggestion to rank constituencies in order of importance and based on their role in the program. What do

Final ACERA Meeting Minutes, 11/5/09

you want to have happen as a result of your communications? Different aspects of program may need to be emphasized to different audiences.

Chair believes White House wants to see all government agencies delivering tangible benefits to the public at large. Audiences have expanded greatly as a result of advanced technology. It was pointed out that the primary need is to first fulfill the mission for which you have been funded before being too concerned with satisfying the public, and also expressing to Congress that there will be a need for ongoing and continuing funding. System is never really finished.

Have to be able to communicate clearly about the distinction between what the ERA is founded to do and the creation of digital cultural content.

Laura pointed out that despite what the committee believes, industry may not come to ERA, committee may need to push the content out.

There was a converse point of view expressed where industry agreement with Generations Network and Footnote.com where industry providers did the digitization. Feeling is that private industry will purchase data, such as Ancestry.com and the 1940 Census data, and ship it out to be name indexed, and the information will be accessed through Ancestry, not the National Archives.

The Library of Congress can not, by law, charge for access to these records so they are not as likely to directly push to promote access to their information to the public. Discussion of utilizing social networking sites -- Flickr, YouTube, et cetera -- to promote access to this information.

Ken commented on previous session that they seemed to be separate from ERA and that involving researchers in determining priorities would add another unnecessary complication to the process. Response was that the record preserver or archivists would need to be consulted as the experts in the field to determine the optimal preservation of each discrete category of materials, i.e., documents, photographs, et cetera.

Chair expressed that the distinction between traditional records and electronic records is a transitional state of affairs, eventually all records will be digital. The Chief Technology Officer of the GPO talked about the GPO model. Their model is 98 percent of publications coming out of the Federal government are electronic. They host those records in State libraries or universities or anybody that wants those links to their site. GPO's special relationship to NARA has to do with National Archives being a Federal repository library. Everything printed by public printer, by law, must go to National Archives.

Final ACERA Meeting Minutes, 11/5/09

Central goal of NARA's Electronic Records Preservation Program is to be able to produce authentic versions of electronic records, and commenter believes authentication is going to become a bigger and bigger issue due to electronic record being inherently not authentic due to the change from its original form. Resource constraints will dictate that original, crumbling piece of paper would be imaged for a fraction of the cost of preserving the original, and original would be discarded, even though that is the true authentic record. Chain of custody is required for true authentication, and that has to be validated prior to ingest to the system.

Question for Lee on this topic of a market plan is how do you see this benefiting the project, what is it you think your need that you're not getting that a marketing plan or exercise might get you?

Basically the message is ERA equals NARA, and actually ERA is going to grow -- NARA as a paper archive is going to shrink over time, and as an electronic records archive is going to grow over time.

NARA can send the message out that this project didn't spend millions of dollars with no results -- something has actually been accomplished here. Have an opportunity to build a marketing campaign with stakeholders -- other cooperative ventures, universities and archivists -- and show the benefit to them to push back a lot of the criticism. From a communications perspective, SAA and ARMA and AIM are some of the professional associations that are completely untapped in terms of marketing.

Andy said you have to be careful about rolling out the message in advance of the implementation, starting to experience some of that with the image interchange framework in that it's been in-progress for 4.5 years and private industry wants to take advantage of progress made so far, but the project is not completed yet.

Think about presented ERA as a storyboard that continues moving and progressing.

Chair concluded the meeting on this note.

6. Adjournment

The meeting adjourned at 11:45