Advisory Committee for Electronic Records Archives
Meeting No. 10
National Archives Building
Draft Minutes
Day 2 of 2, April 8, 2010

In accordance with the provisions of Public Law 92-463, the meeting was open to the public from 9:00 a.m. until 3:00 p.m.

Comments from the Chair

Mr. Ferriero made opening remarks. He summarized a meeting with Vivek Kundra, Chief Information Officer of the United States, this morning, opening a new dialogue and relationship between the National Archives and the Office of the President.

Dr. Robert Martin opened with a reminder that potential actions were documented yesterday. The group will have the opportunity to review the list and determine what they would like to carry forward as action items for the committee. He also asked that action items arising from Committee discussion be explicitly identified as such.

Preservation Prototype Demonstration, David Kepley and Lockheed Martin, Inc.

David Kepley began with a review of where the preservation process was last November. The project started with many ideas and approaches and now they were trying to go from ethereal concepts to practical solutions. The result at this stage is the NARA Preservation Conceptual Framework, which was presented to ACERA last November. They then reached out to the international digital archival community to get their feedback as well.

The current stage has been focused on a working prototype with Lockheed Martin - taking the concepts and putting them into action. This prototype is focused on a framework that will allow the capability to swap out tools and procedures in a flexible fashion. This framework will allow for ingest, characterization, and transformation. The prototype will ensure that the transformations are authentic, and that they leverage the work of other digital archival institutions.

What is Out of Scope - Mr. Kepley made it clear about what they are NOT trying to accomplish in the prototype. First, they are not attempting to determine the optimal format for a given format, but to focus rather on having the ability to transform. Also, the prototype does not include Increment 3 ingest. This is something that is being built into the main system, but wasn't included in the prototype. Additionally, they are not building preservation planning or assessment of risk. Finally, the prototype was not built for scalability, but the framework will be.

So, What's in and What's out? Outside of the scope of ERA currently are the assessment of risk and policies. Inside of the scope of Increment 4 includes Transformation, Characterization, Storage, Workflow of the transformation, and reports.

In determining what was going be transformed, one reason for choosing the formats is because for EBCDIC, there are many files available, and for others because they were free (MS Word and MS Excel)

Mr. Kepley continued by describing the success measurements of the prototype.

1) Filter on records to be transformed - records that have not already been transformed

2) Select a transformation strategy and apply it to any series or transfer group

3) Plug transformation tools into the framework and configure them

4) Transform records to a target format

5) Conduct a trial and archival run

6) Verify the results in reports and by looking at them

7) Verifying the metadata with the originals

An introduction to the Prototype Demonstration was given by Lockheed Martin (Sean Murphy and Kurt Zarefoss)

Some initial questions were addressed about the prototype prior to the demonstration, particularly around the identification of the documents and formats that are most at risk, and developing the system around those formats, as opposed to developing the system to be able to accommodate formats that are not yet known.

Screenshots of the demonstration were given so the committee members could follow along. These screenshots are available as part of the online archive.

The Prototype Demonstration proceeded, covering the following topics:

Demo Page 1
Transformation Strategy
Creation of a Transformation Strategy
Validation Rules
Selecting the New Strategy
Preservation Search
Selecting the Strategy from the Search
Search Results
Initiating Transformation Processing

ACERA Minutes, April 8, 2010

There is a break in the demonstration to first clarify that the original bitstream is kept in the event that a mistake was made and the original file format was needed. Although, when one goes from the original bitstream to a transformation, one is doing this mainly because the original cannot be rendered by a researcher. Another reason is because, as an example, someone may want to convert from an older format, like Wordstar to OpenDocument format or Word format, and the changes that are made in the transformation may not be a change that one cares about. The system will warn you of possible changes. If ultimately one doesn't trust the transformation, then the original is there to use as one sees fit.

The committee refocused the conversation on the prototype demonstration and continued with the job summary reports available in the prototype.

There was a discussion around hyperlinks within the documents and whether it was considered significant to maintain those links, or just record as metadata where those links point to, either inside or outside of the document.

A 10 minute break is taken.

Preservation Prototype Demonstration (Continuation and Summary)

The prototype demonstration continued by completing the review of the job summary log.   Mr. Kepley then provided a lessons- learned summary of the prototype to wrap up the demonstration. For metadata extraction, some of the data desired cannot be extracted with the given tools. There is information gathered on what constitutes a successful transformation,

such as no corruption, maintaining metadata from old to new, and retaining significant properties.

Lessons Learned

Mr. Kepley concluded the presentation by reiterating that NARA already has transfer standards. The reality, though, is that they aren't being paid attention to. A commission, for instance, will be on the backend of their work, and then realize that they weren't following the standards, and then it's too late, because they are already wrapping up, and ultimately place the burden back on NARA.

Finally, Mr. Kepley discussed next steps for the Preservation Prototype.

Next Steps

The Committee turned to a discussion of the Preservation Prototype. The discussion focused on the following topics:

- Significant properties built into the system
- Intellectual property rights to the system and processes
- Barriers to use of open source tools
- Standards of risk assessment
- Costs of transformations vs. maintaining original bitstream
- Challenges of format proliferation
- Common transformation error probability
- Establishing policies, procedures, and standards that require files be "born archival"
- Push the responsibility for the data transformation layer to the end user
- Require that transformations occur earlier in the lifecycle of the record

A request was made for Version 1.1 and 2.0 of the Preservation Framework to be made available to the committee.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

Laura Campbell discussed the Blue Ribbon Task Force Report and Symposium - It was an invitational meeting held at the Fairmont Hotel. A task force was created to look broadly at industry, players, and jurisdictions to put a focus on the economics. A copy of the report is online. It raised the national conversation around the economics of digital content. It enforces the convergence of the challenge of preservation and the economics of public policy.

The committee recessed for lunch, to reconvene at 12:45 PM.

Congressional Testimony Discussion

Malamud Testimony

Mr. Stang initiated this discussion at the request of the committee. Committee members have a desire to speak up about ERA and to be more active in discussions about ERA – to be public advocates for ERA and NARA.   Dr. Martin recalled a working group at a previous meeting of the committee expressing a desire to help out, and he is sure that willingness of the members to assist is still there.

Discussion continued around embracing those that have an opinion about the effectiveness of NARA as partners in improvement.   Hearings are welcome opportunities to get a positive message out about ERA and the impact and importance of having this history for generations to come.

There is a recommendation to complete a case study on ERA that places a context around it. ERA is not just a system. It is the transformation of an environment, a culture, a way of doing work, a new business. To just treat it as a system causes the entire story to be lost. The case study should take the mission of the ERA to the root mission of the Archives.

Also, it's not too late to engage the State Archivists with more communication about what is happening with ERA, focus on bringing State Archivists to meetings, need joint publications that show the council of state archivists and NARA addressing issues together.

Shared Service piece is very important. That's how we got into the state archival business, with blessing from NARA that it would be good idea, and had the backing of Congress. If one has a repository that is trusted and works well, if there's a business model way to provide services, then it is a good idea.

Within the committee, there is an idea to focus on communications. Perhaps take two or three of the committee members and come forward with some ideas - archival, library, general - use the talent you have in the committee that is connected to the industry. Establish a subcommittee that can work directly with NARA. Start by soliciting volunteers.

The Committee moves into a Discussions/Round Table Session.

Mr. David Ferriero - Comments

It was stated that there will be a next meeting of the committee scheduled.

The message around communication has been strong, and the committee will soon see action on that front.

Mr. Ferriero is committed to working with professionals on telling the story about Electronic Records. He believes a documentary is a good idea.

A statement is made that there will be a couple of congressional hearings, and an opportunity for the committee to be called upon to testify.

Dr. Martin stated that he will draft up action items and make sure they are organized into major headings to share with everyone

Mr. Ferriero made a statement about the Open Government Plans and to please look at White House website to read through these plans.

Other Next Steps and Wrap-Up

There was a motion to adjourn moved by Jerry Handfield, seconded by David Carmicheal.

Ayes: Mr. David Ferriero; Dr. Robert Martin; Dr. Kelly Woestman; Jerry Handfield; David Rencher; David Carmicheal; Daniel Pitti; John T. Phillips; Andy Maltz; Dr. Richard Fennell; Laura Campbell; Robert Horton; Dr. Ken Thibodeau; Martha Morphy; Lee Stang
Absent: Dr. Christopher Greer

The meeting is adjourned at 2:56 pm.