

Advisory Committee on the Electronic Records Archives
Meeting No. 11
National Archives Building
Minutes
November 3 and 4, 2010

Welcome Remarks - David S. Ferriero, Archivist of the United States

Mr. Ferriero opened the meeting and expressed his hope that the meeting would reflect the increased preparation the staff put into the meeting and substantive issues placed before the committee for its deliberation. Committee members present: Dr. Robert Martin, Chair, Laura E. Campbell, David Carmicheal, Dennis Day, Dr. Richard Fennell, Dr. Christopher Greer, Jerry Handfield, Dr. Robert E. Kahn, John T. Phillips, Daniel V. Pitti, Jonathan Redgrave, Dan Reed, David Rencher, and Dr. Kelly A. Woestman.

Mr. Ferriero reviewed recent events at NARA:

- The NARA transformation plan, A Charter for Change, was recently released and he asked the committee to review the plan and provide him with comments and suggestions.
- The final year of the Lockheed Martin contract is currently being negotiated.
- The online public access prototype will become available in December.
- Mr. Ferriero and Vivek Kundra, the Chief Information Officer of the United States, hosted a joint meeting of the CIO Council and the Records Managers Council.

Mr. Ferriero introduced Mr. Kundra and invited him to share his thoughts with the committee.

Comments from the Chief Information Officer of the United States - Vivek Kundra

Mr. Kundra outlined the TechStat sessions that the administration has instituted to help manage large federal IT projects. The TechStat sessions convene the decision makers, White House policy office staff, Office of Management and Budget staff, and agency staff, to review and implement large IT projects. After 35 TechStat sessions, \$2 billion has been cut from projects and the deliverables of 19 projects have been accelerated.

Mr. Kundra and Mr. Ferriero have agreed that ERA is one of the high priority projects that the government needs to focus attention on as more of the records the government produces are digital from conception to public release, never existing on paper. The CIO Council and the Records Management Council is working to create a forum to raise awareness of the need to adopt ERA across all government agencies.

Mr. Kundra entertained questions from the committee. Mr. Carmichael, Dr. Kahn and Mr. Pitti were interested in Mr. Kundra's thoughts on the need to preserve platforms and software in

addition to official documents. Mr. Kundra replied that there is a tension between establishing standards for preservation and actually preserving records so that the standard can be applied to them. Mr. Kundra favored establishing criteria for preserving a record to provide guidance to the agencies. Mr. Kundra revealed his ideal in which NARA operated a cloud to which the agencies would be connected. The agencies would have established business practices determining which records would be sent to the NARA cloud and which would not.

Comments from the Chair - Dr. Robert Martin

Dr. Martin thanked the staff for setting up the committee wiki to increase communication between the committee and staff and for providing briefing materials on the wiki in advance of the meeting.

Adoption of Minutes (April 7-8, 2010)

Dr. Woestman moved adoption of the minutes of the April 7-8, 2010 Committee Meeting. Mr. Pitti seconded the motion. The motion passed with 13 in favor, none opposed, and Dr. Kahn abstaining not having attended the April meeting.

Review of Action Items

The committee reviewed action items that have not been completed.

Action Item 11 - "explore a broad discussion around data models and metadata and develop a white paper on the best practices for records metadata." This item will be picked up in committee discussion.

Action Item 13 - will be considered by the committee during the meeting.

Action Item 14 - engaging state archivists. The committee extensively discussed the possibility of scaling ERA so that it could be exploited by state archivists. The NARA contract with Lockheed Martin for ERA implementation limits use of ERA to the federal government. Mr. Reed suggested the states collaboratively develop an open architecture system to take advantage of the NARA APIs since the software development task will require a huge effort. Consideration of Action Item 14 will continue later in the meeting.

Action Item 15 - prepare a report on meeting strategies. Mr. Rencher prepared a draft report on meeting strategies and it has been circulated for comments. Committee action on this item was deferred to after the presentation by David Lake and incorporated into the Facilitate Discussion item.

Action Item 16 - assist NARA in thinking about what's next. This item has a slot on the meeting agenda.

Action Item 20 will be considered with Item 11.

Action Item 21 - distribute risk assessment and archive-ability standards to all agencies. In process.

The committee recessed at 10:13 a.m. to 10:28 a.m.

ERA Update - Meg Phillips and David Lake

ERA currently has under management around 86 terabytes of data and around 314 million objects. The vast majority of that total comes from the Bush White House.

Two prototypes were completed this summer, one for online public access and the transformation prototype, which handles format migrations.

Ms. Phillips has been appointed electronic records life cycle coordinator, a new position. Her responsibility is to make sure that all the different offices within NARA which are affected by the ERA project have a seat at the table and a voice in what is built

ERA Business Requirements Group met a number of times to set the high-level goals to be built into Lockheed Martin contract for 2011.

- A system that federal agencies can use for their interactions with the National Archives, agencies must be able to schedule records with NARA in any format and arrange for the accessioning of records in any format
- TechStat session on ERA reduced by six months the amount of time NARA has to develop ERA, development must be completed by the end of fiscal year '11
- Lay the groundwork for format migration in the future
- Launch the online public access interface
- Prepare the EOP system that supports the Bush records right now for the 2014 opening of those records to FOIA requests.

Recent negotiations with Lockheed Martin have resulted in very realistic options for what can be accomplished in the final year of the contract. The schedule rather than the budget became the limiting factor for what could be accomplished in the final year of the contract. NARA reduced the scope of what it requested to ensure Lockheed could deliver a completed product rather than an incomplete broader product. Capabilities NARA desires but had to remove from the scope of work include:

- data type template verification
- a more flexible public assess architecture than Vivisimo provides on its own, which would make it easier to incorporate searches of additional data sources

NARA goals sacrificed by decreasing the scope of work in the contract:

- The four legacy systems that support the National Archive electronic record process will not all be shut down as early as planned
- NARA will not be able to demonstrate sophisticated preservation planning in ERA, though it will be done and documented outside ERA.
- NARA will not be able to make needed improvements in life cycle management of records from the time of agency submission of a new record schedule forward inside ERA.

Negotiations with Lockheed are expected to conclude by November 10, 2010. Ms. Colon added that NARA has learned that it is better to receive many small finished products from Lockheed than a few big products; Lockheed is more productive and NARA is better able to absorb the product and provide feedback to Lockheed.

NARA deployed ERA in a pilot phase to get feedback from a small group of agencies. Feedback from the agencies included: this is hard, this is confusing, the training's not quite good enough, and the user account request process is difficult. NARA will poll agencies over the remainder of the year asking them to commit to a time when they will start using ERA. NARA hopes to move a few agencies a month to ERA from July 2011 through the end of the 2012 fiscal year.

Committee comments for Ms. Phillips:

Mr. Reed observed that standard lines of code is an archaic method of measuring level of effort in software development and of limited value for measuring level of effort. Dr. Kahn asked if NARA has considered a mechanism whereby the authentication of the record can be developed from the original source and maintained forward. Ms. Phillips answered that NARA supplies agencies with a wrapping tool that applies a hash value to a set of records which can then be used for authentication and validation

David Lake presented issues he is addressing in the ERA communications plan. These included the roll out of online public access in December 2010, managing the expectation of users, training agency staff as ERA gets rolled out to more agencies in the coming months. He also offered to provide a stock presentation to committee members and other stakeholders can use when they speak about ERA.

Committee comments for Mr. Lake:

Dr. Martin said the archival community represented by SAA, CoSA, and NAGARA should be given greater emphasis in the outreach to stakeholders. Mr. Phillips urged NARA to keep communications a two-way flow between NARA and the stakeholder community. Mr. Phillips also urged NARA to request and address feedback from the community so they feel engaged in the value added by ERA. Mr. Reed and Dr. Rencher said Mr. Lake should produce stock

presentations for the different sets of stakeholders NARA needs to reach, i.e., government agencies, technical audiences, the archival community, and the general user.

Office of Management & Budget (OMB) Technical Status (TechStat) Review - Charles Piercy and Laurence Brewer

NARA underwent a TechStat review on June 17 when OMB was concentrating on programs that appeared to be in trouble. OMB went on to refine the TechStat review process to focus on high-priority programs and NARA received a second TechStat review on August 10. ERA was the primary subject of these reviews. A follow-on TechStat review is scheduled for November 9. Direction from the first two TechStat reviews include:

- Get the archival community galvanized around ERA
- Instill confidence in the ERA system by opening lines of communications and showing that NARA had made objective progress.
- Put together a plan to implement the highest priority requirements that NARA needs to make the system operational

The TechStat reviews identified issues needing corrective action:

- A lack of detailed plans for the final two increments.
- Low usage of the system
- Improved strategic planning, including input from external stakeholders

Actions taken by NARA to address TechStat concerns:

- Committed to the goal of ingesting 10 terabytes per quarter for a total of 40 terabytes per year into the archives
- Produced an ERA implementation plan by identifying which prioritized requirements can be completed in fiscal 2011, ensuring that ERA's system is able to meet its initial operational goals.
- Plans to award a firm fixed price contract to support O&M beginning October 1st, 2011.
- Convened a meeting of ERA external stakeholders as part of the SAA/CoSA/NAGARA Conference aimed primarily at state archivists.
- On October 20th convened a government stakeholder meeting, the first joint meeting of the Federal Records Council with all of the records officers from all the federal agencies, and their corresponding CIOs.
- Developed a road map for a strategic plan. NARA has not created the strategic plan because the needed inputs are not yet ready.
- Operations and management-only contract - NARA plans to take a couple of years off on development, determine a strategic approach and then try to get a budget initiative for 2014 and then resume development in 2014.

Committee comments for Mr. Piercy:

Mr. Piercy described Mr. Kundra's hope that NARA store and manage temporary records in a cloud and said NARA pushed back on that idea not feeling ready to dive into such a vast undertaking. Mr. Carmichael, Mr. Pitti and Mr. Phillips encouraged NARA to enter this territory now while the standards and practices are being developed so that NARA can provide input and guidance into this development and not be put into the position of having to retrofit a system to standards that may not suit NARA's vision and capabilities.

Mr. Phillips urged NARA to abandon using the term "electronic record" and to use "digital record" instead.

Facilitated Discussion about OMB TechStat Review - Meg Phillips

The committee discussed how it could be an asset to NARA, many members commented that they felt their previous work on the committee had not helped NARA and that the staff time consumed to provide the logistics for committee activities may have been a greater minus than any plus the committee's activities provided NARA. The committee discussion consumed the afternoon of the 3rd and morning of the 4th.

Mr. Redgrave proposed a structure for committee discussion and future work. He presented a diagram with four headings: Legal Policy Construct, Business Process, System, and ERA Core. Mr. Redgrave explained his conception of these topic areas and the potential scope of committee activity in each area. The committee adopted his proposal as a structure for discussion and concluded with five topic areas for subcommittee organization: 1) Legal Status, 2) Technological Issues, 3) User Interface, 4) Development Strategy, and 5) Communications.

Ms. Phillips asked the committee to concentrate on the middle and long term issues since the time frame for negotiating the final year of the Lockheed Contract is too short for the committee to have a meaningful impact.

The topics the subcommittees work on will overlap, it is expected that two or more subcommittees will address different facets of the same topic area. The committee agreed that the work of the subcommittees will be updated through the wiki and teleconferences. Committee members should feel free to participate on different subcommittees as their interests and time allow, but subcommittee leads will not abandon their topic.

Ms. Phillips said that NARA staff will be wrestling with many of the same issues as these subcommittees. If staff monitors the wiki there may be occasions when staff would invite committee members to participate in conference calls to resolve a particular issue

Subcommittee 1 - Legal and Policy Construct

Jonathan Redgrave, lead, Bob Kahn, Jerry Handfield, Daniel Pitti, Bob Martin.

Review the statutes applying to NARA and how they should be revised. NARA works with concepts of "record" and "archive" derived from statutes that were promulgated before the digital

age. Congress will have to update these statutory definitions at some time and the committee may be able to help NARA evaluate how these concepts are defined from an archival perspective as well as from a records management and compliance perspective. If this reevaluation includes how agencies produce records the budget implications become much larger than NARA alone.

The federal government needs to develop a job category for record manager. If the federal agencies are going to store records in a cloud the process by which they do so needs to be standardized if the records are going to have value.

Other issues the subcommittee may address:

- Creating a unique identifier for digital objects is one the fundamental tasks needed for the system to succeed.
- Revise the real and statutory relationship between NARA and GSA.

Subcommittee 2 - Technological Issues

Bob Kahn, lead, Dennis Day, Laura Campbell, Daniel Pitti

Proposed White Paper, What is a Record? The Long Term View

What is the definition of a record and what kinds of metadata schemes might apply to non-documentary records such as videos, spreadsheet, graphic presentations, and chemical structures the future will present.

Open architecture may become a subsequent topic of focus.

Subcommittee 3 - User Interface

Kelly Woestman lead, David Rencher.

Look at user expectations outside the archival community. What are the needs of students, teachers, and the general public. The Google interface is a standard expectation. Find out what works for ERA users who are not archival professionals. As mobile devices become more ubiquitous sources of information, access via the web browser and PC will become less prevalent.

Subcommittee 4 - Development Strategy

John Phillips, lead, Daniel Pitti, Richard Fennell

Many members of the committee argued for parallel development of an open architecture system. The federal government will have to take a lead in developing a new system and establishing standards; states and private industry are wrestling with the same issues. If the federal government devises a workable standard, the other sectors may fall in line behind it because a) the federal government will have done most of the work and b) having an established federal system will provide cover for any short comings the other sectors may encounter. The states would have an incentive to collaborate in developing such a system which they then could then

exploit. Such an open architecture system would entail a perpetual development regime such that after 5 years the ERA system will have evolved to an extent that very little of the 2011 code will still be in place in 2016.

Ms. Campbell cautioned that no matter what model is used, getting the metadata right and the definition of the record right is of paramount importance. Mr. Pitti suggested adopting the term "control data" to replace "metadata" as a more accurate description of the data tag.

Records management and compliance needs to be included in software procurement contracts so the track for record preservation and management is "baked in" from the moment the record is created. As the federal government moves from the creation of millions of records during one administration to billions and inconceivably larger numbers in the future searching and finding a record of interest will depend on good metadata being appended to every record, the creation of good metadata must be automated.

Subcommittee 5 - Communications

David Rencher lead, Dennis Day, and Jerry Handfield.

Work with David Lake on the appropriate message NARA should have prepared for each of its stakeholders.

Online Public Access Update, Feedback & Demo - Pamela Wright

Online public access to ERA will go online in December. It employs a user-centered approach to encourage citizen engagement. NARA wants the public to know what records are available on a particular topic and what the archival context is for those records. The vision for online public access is to have the online catalog, the electronic records, everything on archives.gov, and the Presidential library Web sites all searchable through one interface.

NARA has compiled user feedback. Users love the wealth of information online, and that it is free and convenient. They want everything to be put online and they want NARA to be clear about what is online and what is not online.

NARA has moved into social networks over the last year. Ms. Wright demonstrated the IndyCommons app on an Apple iPad. The IndyCommons was put together by the Smithsonian and Archives participated. Through Flickr NARA photos along with metadata are publicly available. This was accomplished at no cost to NARA.

Presidential library Web sites will not be included in the prototype rolled out in December. The OPA display of archival hierarchy will hopefully make more sense to the general public and be used more often than it is today. Search results are organized to look like Google search results because that is what the public is comfortable with and has asked for. Ms. Wright admitted the terminology used in OPA is still archival in flavor and hoped feedback on the prototype will provide more user friendly language.

The pitch to the public during the rollout will be a splash page with the message "Coming soon: Online public access search and display interface. Give it a try and give us your feedback." NARA plans to add image zoom capability before 2012. Continuing the effort to increase citizen engagement NARA plans to add the capability for users to tag information online by May 2012.

Committee comments for Ms. Wright:

Mr. Phillips suggested NARA find a way to have stakeholders contribute and fund record sets as a way of increasing stakeholder engagement. Dr. Kahn said NARA should put it in its long range plans opening up its metadata architecture so private entities could create their own Archives app, tailored to their markets or interests.

Update on Transformation Prototype (Preservation) - David Kepley

NARA is considering releasing the transformation prototype in increments as individual elements of it are developed to get feedback from the field, and asked committee members to help alert stakeholders of the opportunity to peruse the prototype and comment on it as the releases are made public.

NARA is exploring initiatives to ensure documents destined to become permanent records are captured when they are created. Major impediments to preservation are the loss of metadata, if a record's metadata is lost the record becomes meaningless, and the issues of hardware and format obsolescence.

NARA can provide access to permanent electronic records by letting users download records, providing viewers the public can use, or NARA will transform records into a format the public can use. That transformation into a usable format is a challenge NARA is currently addressing.

For obsolete formats, such as EBCDIC, NARA will prepare white papers addressing the technical issues of transforming the format to ASCII or some other standardized format. After the technical white paper is written and circulated NARA will produce a preservation plan and an access plan for the format. These plans will be released to the public for comment.

Committee comments for Mr. Kepley:

Mr. Carmichael urged NARA to share the transformation challenges on a wiki so that intelligent input can be garnered from all over the world, because NARA is not the only institution wrestling with transformation issues. Mr. Phillips advised NARA to alert the agencies about the need to submit records in usable formats because NARA does not now and will not in the future have the resources to transform any format. Some of the onus of usability has to be shifted off of NARA.

Dr. Kahn suggested NARA make the transformation tools available to the public on a NARA website akin to an app store. If format converters could be downloaded by users then

conversions of formats could take place on the users' computer after they downloaded the data file they were interested in.

New Action Item Review and Wrap-Up

Dr. Martin said the newly constituted subcommittees should post progress reports on the committee wiki by January 31, 2011. The next committee meeting was scheduled to occur in the first or second week of April. Dr. Martin adjourned the meeting at 3:10 p.m.