
Progress
Report:
September
1, 2008-
February
28, 2009

NHPRC EMCAP Tool

Kelly Eubank



NORTH CAROLINA DEPARTMENT OF CULTURAL RESOURCES

BEVERLY EAVES PERDUE, GOVERNOR
LINDA A. CARLISLE, SECRETARY

OFFICE OF ARCHIVES AND HISTORY
JEFFREY J. CROW, DEPUTY SECRETARY

TO: Lucy Barber, National Historical Publications and Records Commission
FROM: Kelly Eubank, Project Director, Electronic Records Archivist

SUBJECT: Electronic Mail Preservation Collaboration Initiative, Grant Number NAR RE-05701-07 (North Carolina, Kentucky, and Pennsylvania)

Progress Report: September 2008-February 2009

The Electronic Mail Preservation Collaboration Initiative aims to build a prototype tool for the collection and preservation of archival electronic mail. Staff from the states of North Carolina, Kentucky, and Pennsylvania have teamed together to leverage their knowledge and skill sets to help design and test a tool that will collect electronic mail, maintain a copy of the original email and the attachment, as well as transform the email itself from its native format to an XML format based on an XML schema. The team developed on-line training materials. The project team includes: Kelly Eubank (NC), project director, Druscilla Simpson (NC), head, IT Branch; David Minor (NC), programmer; Ed Southern (NC), head, Government Records Branch; Mark Myers (KY), electronic records archivist; Glen McAninch (KY), head, Technology Analysis and Support Branch; Linda Avetta (PA), information technology generalist administrator; Cynthia Bendroth (PA), head, Division of Records Administration and Image Services.

MAILING ADDRESS
4610 MAIL SERVICE CENTER
RALEIGH, NC 27699-4610

TELEPHONE 919-807-7280
FAX 919-733-8807

LOCATION
109 EAST JONES STREET
RALEIGH, NC 27601

Information Technology Environment:

Each of the three partner states has, or is in the process of consolidating into, a centralized IT environment. In view of this fact, the Electronic Mail Collection and Preservation (EMCAP) team held a conference call in September to discuss whether hmail might be synchronized to Active Directory (AD) in Microsoft Exchange. Kentucky began discussions with their central IT group about the feasibility of this step and about rolling this out to a slightly wider audience. The programmer suggested using a Secure Socket Layer (SSL) server to authenticate the EMCAP server through to each of the state identification systems, rather than syncing directly to AD. From a security standpoint, this creates fewer potential problems than a direct sync to AD. From the perspective of wider use, using a Secure Socket Layer makes the tool more open and able to be used by other institutions. If we choose to go the route of syncing directly to AD, it makes the tools proprietary and useful exclusively to one e-mail platform, Microsoft Exchange.

In North Carolina, many of the state's IT resources continue to be consolidated under Information Technology Services. In order to complete this consolidation in a timely manner, staff from the Department of Cultural Resources/Information Technology group (DCR-IT) dedicated part of their time to assist in this effort. As a result, many of the internal projects in which they were involved were put on hold. David Minor, the programmer on the EMCAP project, was pulled off many of his projects in order to help with the consolidation. Additionally, he had to take on management duties because his supervisor was out for several weeks on medical leave. Having involved in e-mail preservation since 2005, however, David continued to do some programming in off hours to ensure that this part of the project was completed. He reviewed the code in detail that the temporary staff person wrote earlier, cleaned it up, and tested it. In some cases, he chose to begin from scratch as the code did not prove serviceable. In addition, he added some features to the tool that were not in the original specifications.

The EMCAP group continued to collaborate with the Collaborative Electronic Records Project (CERP) group on refining and testing the XML schema. The schema is based on the RFC 2822 standard for email structure. According to the programmer on the CERP project, using the schema, CERP was able to parse successfully ninety-nine percent of the e-mail they collected. Older e-mail, created before the RF2822 protocol, has proven to be a challenge.

In January, as the deadline for completion of the software came and went, the NC group met to discuss the grant and software status. The team explained to management that while David worked on the project, it was not his priority as originally planned. Once his supervisor returned to work, David was relieved of those duties and committed to spending thirty two hours a week on the project. One of the key points during this time was to discuss what still needed to be completed, what part of the original work plan had not been completed, and what could be cut from the programming items. The group decided that the "exporting of the files to .pst files" could be eliminated, as it was not part of the original grant.

In the January conference call, the group discussed our respective IT environments for the grant and whether or not it made sense to ask for an extension or simply end the grant at the end of February. Despite differing levels of opportunities within our respective IT environments, we all agreed to seek an extension. The group also discussed the remaining \$11,000 that had not been spent from the grant. Each state was in favor of using the money to move some of the

tools out of their proprietary formats and into an open format. We agreed to forego the closing meeting.

State Partners

Emergency Management (NC)—We continued to try to work with Emergency Management. Unfortunately, the timing of this work was not advantageous. The approach of the hurricane season, which begins in June and lasts the end of November, required Emergency Management to focus on storm preparation. NC was able to load a copy of Lotus Notes from Emergency Management onto the laptop to begin using and developing training tools. However, Lotus Notes does not allow users to copy emails. Megan worked with Pete Farmer, head of IT support for Emergency Management, to address this issue. However, because of their busy schedule and Megan's departure, NC still has not been able to get Emergency Management onto the hmail server.

Governor's Office (NC)—the head of the Citizens Services and Community Affairs office in the Governor's Office continued to successfully transfer constituent e-mail using EMCAP. In January, the Governor's office changed leadership in January. Governor Beverly Perdue took the oath of office in January. The ER unit concentrated on helping with the transition of Governor's records from Governor Easley. Claire Ennis, formerly with the Governor's Office of Citizens Services and Community Affairs left her position as head of this office to take a position with the State Controller's office. In February, Kelly Eubank and Pam Ingle met with staff from Governor Perdue's office of Constituent Services. The Governor has not appointed someone to lead this office but the staff still with this office agreed to continue using and testing EMCAP.

As of February, there are 39,071 e-mails from Citizen Services and Community Affairs.

Secretary of State (NC)—The Secretary of State's office has not deposited emails into EMCAP since the first transfer. Kelly contacted them in January to encourage them to deposit email into EMCAP. The reporting function in EMCAP has not been completed, so it is difficult to tell if further deposits have been made.

Training

Before she left at the end of October 2008, Megan cleaned up the EMCAP Web site and posted meeting minutes from 12 previous meetings to the Web to it October. These minutes were from meetings over the course of the grant.

Pennsylvania and Kentucky:

Pennsylvania

Pennsylvania updated the user documentation for specific use in PA and trained ten executive staff individually during the months of August and September. To assist users and make it easier for them to use EMCAP, folders were created for each individual based on the Commonwealth General Administrative Retention and Disposition Schedule. The records most commonly used from the General Schedule were also attached at the end of the user manual. We held a meeting in December with the staff involved and found a variety of uses. Some hadn't used it at all. Others actively used it the first month or two but then slacked off and were either no longer using it or felt they didn't have anything that was archival. One individual used the tool inconsistently. We discussed with them the reasons they were no longer using it and the majority said it was due to lack of time. [Just a parenthetical comment: this will be a common complaint with this or any similar tool, which is probably stating the obvious.]

In addition, the Pennsylvania team along with the Office of General Counsel and the Office of Enterprise Records Management provided e-mail-specific training during the months of September through December for the new Right-To-Know-Law that took effect January 1, 2009. The awareness alone was worth the training.

Kentucky

In September an IMAP connection was established from retired commissioner Jim Nelson's account (being reviewed by records analyst Jim Cundy) to the EMCAP server. Review of the over 6,000 e-mails (over a 10 year period) in the account continued throughout the report period.

Discussions with the Office of the Secretary of the Education Cabinet continued during the period with the focus on identifying a cabinet specific e-mail policy and training staff in management of e-mail in order to later implement an IMAP connection to the EMCAP server. Training in e-mail management with other state and local agencies also continued.

In September Glen McAninch and Skip Hunt met with the head of the Commonwealth Office of Technology E-mail working group. The EMCAP grant and the network architecture of the server connections including Active Directory were explained and the COT representative was satisfied that the project configuration was fully compliant with Enterprise Architecture and COT policies.

Outreach

In October, the NC group met with Alexandria Everleigh of the Winston Churchill Memorial Trust to discuss their efforts with regards to outreach and preservation of electronic records. Alexandria explained the organizational set up of her Archives in the UK and the difficulty getting funding to tackle issues.

Megan updated the EMCAP website with the SAA presentation and XML schema documentation.

November 10, 2008, Kelly Eubank and David Minor attended the "Preserving Historical Correspondence Symposium" in Washington D.C., hosted by the Collaborative Electronic

Records Project as a part of their project wrap up. At the time, David Minor and the programmer from CERP had collaborated on developing the schema. The CERP project, using a parser built in Squeak, a variant of Smalltalk, had successfully parsed 99% of the e-mail they collected. The EMCAP parser had not been built to test this theory, but David had parsed a few of the e-mails we collected and it seemed to parse properly. The symposium was attended by members of the federal government, people at Harvard, and members of the faculty at the University of Maryland.

The group has also received e-mail inquiries about the tool and whether or not it the code can be shared. Once the project code is written, we will determine how best to share it. In September, Kentucky answered an e-mail inquiry about reading archived .pst files from Michael P. Martin of the New York State Archives.

In NC, as a result of the meeting regarding NC's budget appropriation bill, specifically section 6.12, the Archives set up a meeting with the Director of Enterprise Services and the NC Mail Product Manager to discuss steps for using EMCAP with Microsoft Exchange and for moving EMCAP into production in state government.

In December the group had a conference call to discuss with Lisa, a graduate student from Toronto, the possibility of doing a case study related to the EMCAP project. It is likely that NC will be the focus of the study, with KY & PA being points of contrast or comparison.

Potential Partners

In NC, with the changing of the Governor's administration, Claire Ennis left the Governor's Office for the Office of the State Controller. She gave very positive feedback on using EMCAP. She wholeheartedly supports the continued use of the tool, saying that once everything was worked out, it was very easy and efficient to move data.

NC contacted the Secretary of State's office to check in with them on transferring data. However, they have not responded and have not copied e-mails.

In Pennsylvania, as previously noted, it appears that PA may be only able to test within their agency, the Pennsylvania Historical and Museum Commission. Should another opportunity arise, we will certainly pursue it.

Administrative Items:

Megan Durden, Archivist with the Electronic Records Unit in NC, left state employment in October 2008. While her job was advertised, a slowdown in NC's economic state prevented her position from being filled.

Governor Michael Easley, in one of his last acts as governor, on January 9 issued Executive Order 150. EO 150 specifically concerned E-mail and Email Archiving. See addendum.

The EMCAP team applied for and received a no cost extension until June 30, 2009 to complete the programming and testing of the tool.

David Minor and Kelly discussed what would need to be done in order to move the proprietary items into open source. Because of the restrictions in place and the steps required to bring in

additional staff, we decided to forego that. Additionally, it would require much extra work on David's part, work with which his supervisors would have been unhappy. Instead, we decided to use the remaining funding to hire a temporary to help with the documentation of the tool as promised in the original grant submission.

State of North Carolina



MICHAEL F. EASLEY
GOVERNOR

EXECUTIVE ORDER NO. 150 E-MAIL RETENTION AND ARCHIVING

WHEREAS, the North Carolina Public Records Law declares that the public records and information compiled by the agencies of North Carolina government are the property of the people; and

WHEREAS, all e-mail messages sent and received in connection with state business are public records; and,

WHEREAS, a transparent government and the citizens' right to access public records, are of paramount importance; and

WHEREAS, as a result of changing technology and the need to ensure that public records are properly preserved, I established the E-Mail Records Review Panel to review and recommend changes to the current e-mail and electronic text communication record retention policies for North Carolina's executive branch agencies; and

WHEREAS, the E-Mail Records Review Panel met six times, which included public hearings where the Panel heard from representatives from the North Carolina Press Association, the North Carolina Association of Broadcasters, the State Employees Association, and other interested parties; and

WHEREAS, the E-Mail Records Review Panel submitted to me its recommendations and proposed changes to current e-mail and electronic text communication record retention policies; and

WHEREAS, I have carefully reviewed and considered the E-Mail Records Review Panel's recommendations and proposed changes regarding current e-mail and electronic text communication ("e-mail") record retention policies.

NOW, THEREFORE, pursuant to the authority vested in me as Governor by the Constitution and the laws of the State of North Carolina, **IT IS ORDERED THAT:**

1. Executive Branch employees shall treat all e-mail messages, which they send or receive via state government e-mail accounts as public records and shall handle and maintain them in compliance with the Public Records Law and records retention policies in the same manner as paper documents or other tangible records.
2. Executive Branch employees shall not delete in a 24-hour period any e-mail messages sent or received in the course of conducting State business.
3. Executive Branch employees shall not use state e-mail accounts for political purposes.
4. Executive Branch employees shall not use state e-mail accounts for personal purposes except to communicate about family matters. All employees shall assume that information on the state's e-mail system is subject to public review and to review by state officials.
5. All outgoing e-mails sent from Executive Branch state e-mail accounts shall include language notifying the recipient(s) that the message is subject to the Public Records Law and may be disclosed to third parties.
6. Employees who conduct public business via personal e-mail accounts or non-government technology shall ensure that all public records are properly retained and archived pursuant to the Public Records Law and applicable record retention schedules.
7. The Department of Cultural Resources (DCR) shall provide all Executive Branch employees with online training for managing e-mail as public records, which training shall be mandatory for all employees who handle public records.
8. Information Technology Services (ITS) shall copy all Executive Branch agencies' e-mail messages to backup tapes at least once daily and shall retain the tapes for a minimum of ten (10) years.
9. Executive Branch agencies shall collaborate with the State Chief Information Officer (CIO) and DCR to employ a software platform that complies with the E-Mail Records Review Panel's recommendations, including saving backup tapes for a minimum of ten years.
10. As soon as possible, the Office of the State CIO shall procure, through the competitive bidding process, an archive system and shall work jointly and in collaboration with DCR to provide the archives/records management software package to be used by state agencies.

11. Executive Branch agencies shall follow all other directives issued by the Office of the Governor pertaining to e-mail retention and archiving policies, consistent with North Carolina law and record retention schedules.
12. DCR shall conduct random audits of state agencies in the Executive Branch to ensure that employees are in compliance with the records retention and disposition schedules and DCR shall conduct annual reviews of backup tape requests and provide reports to the State CIO and the Office of the Governor.
13. State agencies outside the Executive Branch and not directly subject to this order are invited and encouraged to review and revise their e-mail and electronic text communication record retention policies consistent with this Executive Order.

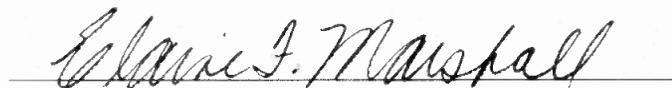
IN WITNESS WHEREOF, I have hereunto signed my name and affixed the Great Seal of the State of North Carolina at the Capitol in the City of Raleigh, this the ninth day of January in the year of our Lord two thousand and nine, and of the Independence of the United States of America the two hundred and thirty-third.





Michael F. Easley
Governor

ATTEST:



Elaine F. Marshall
Secretary of State