

---

Progress  
Report:  
January 1,  
2008-  
February  
29,2008

---

NHPRC EMCAP Tool

---

Kelly Eubank

---



## NORTH CAROLINA DEPARTMENT OF CULTURAL RESOURCES

MICHAEL F. EASLEY, GOVERNOR  
LISBETH C. EVANS, 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: North Carolina, Kentucky, and Pennsylvania's Electronic Mail Preservation Collaboration Initiative, Grant Number NAR RE-05701-07

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 is also developing on-line training materials. In addition, once the tool has been designed and tested, we will deploy this tool to participating pilot partners for phase two testing and feedback. 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.

### **Objective 1: Information Technology Environments**

1. Continue the development of the e-mail collection and preservation software to convert e-mail from its native format into the more stable XML format and complete written documentation for the program.

**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:**

The EMCAP group met with David Gilmore (IT), Bruce Garner (CIO) and Haley Haynes of the Secretary of State's office about their participation in the EMCAP grant. They remain keenly interested in partnering with us and identified the corporations division for this pilot. Mr. Gilmore mentioned that GroupWise is a proprietary system and the group discussed whether or not SOS would be able to participate. Pamela Ingle agreed to follow up with Cheri Myers to discuss the records disposition schedule for the corporations division so that they could participate. Kelly followed up the meeting with a phone call to Mr. Gilmore to discuss the email mechanism and accounts on their virtual private network (VPN).

In NC, the consolidation of section IT staff continued. David Minor now reports to the Applications Development Team under Department of Cultural Resources Information Technology Division. In addition, DCR-IT was charged with consolidating the servers throughout DCR into the newly constructed server room in the Archives and Library building. Because this was such a major undertaking, all staff of DCR-IT was pulled in to work on this effort to smooth the transition. On a broader scale, DCR-IT will be consolidated under the central Information Technology Services, (ITS). All desktop support and network personnel will report to ITS. We are still uncertain as to what will happen to the programming staff.

### **Programmer Update:**

On January 10, the NC team held a meeting with the applications developer and the contractor hired for the project, for an update of their work. The programming team is working from a milestone document that was developed in October 2007. (See addendum A.) The first milestone is to set up the accounts (Administrative Console) Batch account and create a user access website. Part of the work is an establishment of a framework so that all the parts are there. In addition, there will be a secure, web-based utility that a can be used by a user to check the status of his or her account or make a request that his or her password be reset. The approximate timeline for this milestone is November 1-January 16. Because the contractor did not start until Nov. 13, the timeline has some flexibility with regards to the Jan. 16 deliverable day.

In the January 10 meeting, the applications developer and the contractor conducted a demonstration of the capabilities for the first milestone. While there was some clean-up work to do, the console seemed to be working. The applications developer said that they would concentrate on cleaning up the code and making buttons work and be ready to demonstrate to the larger team by February 8, 2008. Kentucky has web-meeting software and desktop-sharing software they can make available to us for this demonstration.

At the beginning of February, the applications developer decided to move the contractor to phase two of the project while he cleaned up the code from phase one. He indicated that phase one would be fully complete by February 22. By February 13, 2008, the applications developer expressed concerns regarding the skill set of the contractor to complete the second milestone. Milestone two is defined as completing development work on the core processing engine and XML transformation.

The XML transformation requires a parser to parse each element within the email and transform each element to XML. The central processing engine should be able to import and export .pst files (large ingest), record individual actions to the account, and compare messages in order to process only the changes made.

On Feb. 28, 2008, the NC team met with our Chief Information Officer and two members of the application development team to discuss the progress of the programming. In addition, the archives staff wanted to discuss what opportunities were available to allow our applications developer to spend more time on this project. Because of consolidation efforts, the applications developer has been pulled into performing networking duties in order to get everything to sync. As a result, he was not able to devote the time he wanted to supervision and development. The CIO indicated that he was willing to work with us in terms of time commitment but stressed that this project is not the only project DCR-IT currently has. In a subsequent follow-up meeting with the contractor, the CIO committed to allowing our applications developer to spend eighty percent of his time on this project and with the contractor to get it back on track. We are to evaluate the progress at the end of February.

By February 20, our applications developer released a working version of administration console (milestone one) and gave Kelly Eubank and Megan Durden administrative rights to that database. These rights allow staff to set up accounts and create passwords for users. The hmail server can collect email once the client is configured to send to our server. However, because the email parser has not been completed, the email collected is not normalized. The applications development team in NC is working closely with the contractor to complete the rest of the work. We will be looking to deploy the custom software to Kentucky and Pennsylvania to load and test.

The partners held a conference call on January 16 to discuss updates on the project and assess progress. The Kentucky team announced they set up a wiki as a way for us to communicate and document the progress of the project. Kentucky is also providing infrastructure to help the project operate. NC does not have this capability. The wiki allows us all to communicate and add documentation and slides. Additionally, in the Kentucky VPN, Skip Hunt loaded SnagIt software to allow us to create screen shots for training and documentation purposes. The Pennsylvania team had some difficulties accessing the KDLA accounts and the hmail accounts. As a result of the call, they were able to access the material.

### **Pennsylvania and Kentucky:**

Linda Avetta submitted the Technical Architecture Review (TAR) document to Pennsylvania's central OIT (Office of Information Technology) group. This document came up for discussion in a February meeting. As a result of that meeting, a conference call with Pennsylvania's OIT group was scheduled for February 29.

### **Objective 2: Potential Partners**

2. Test the software to determine its scalability, efficiency, and employee compliance, using records retention schedules.

### **Potential Partners**

#### **Pennsylvania:**

Pennsylvania would very much like to work with two entities for this project—the Governor's Office / Office of Correspondence and the Pennsylvania Historical and Museum Commission

(PHMC) Executive Office, of which Barbara Franco is the executive director. The internal executives of PMHC have agreed to participate. At this time, the question still outstanding is the issue of using beta software that touches the production environment. Pennsylvania's OIT group required a completed Technical Architecture Review (TAR) document for approval of the beta project. Linda Avetta indicated they submitted their Technical Architecture review to OIT for approval to begin the project as noted in Objective 1 above.

### **Kentucky:**

The staff in Kentucky has the email from the outgoing commissioner of the Kentucky Department of Libraries and Archives (KDLA) as well as the agreement of the current commissioner of the KDLA to participate. An election has been held and a new administration took office in December 2007. The outgoing Secretary of the Education Cabinet determined that the e-mail of the previous Cabinet Secretary contained no archival records. The Archives plan to work with the new administration to establish an archiving procedure for e-mail with Cabinet secretaries and perhaps the new Governor's constituent e-mail. As of the end of February, only preliminary meetings had occurred with new administration officials.

### **North Carolina:**

#### **Secretary of State**

In January, the NC team met with Haley Haynes, Bruce Garner and David Gilmore of the Secretary of State's office to further discuss the project and timeline for deliverables. The NC General Assembly begins their legislative session in May of 2008, so NC would like to deploy the tool in April before then. The group also included Pamela Ingle from the Records Management Group and analyst for the Secretary of State's office. Ms. Haynes identified the Corporations Division as a good candidate for testing this tool. The Secretary of State's office uses GroupWise for their email system. A number of technical issues arose in this meeting. (Please see Consultation Summary, Addendum B.) Mr. Gilmore is the IT person in charge of their system. He is going to review those technical issues and get back to us. Ms. Ingle has been in contact with the head of the Corporations Division and drafted a retention and disposition schedule for their approval. After talking with that Division Head, we determined that only a handful of people would be testing the tool. Ms. Haynes agreed that she would like to keep it to a small number for now.

#### **Emergency Management**

After the meeting with the Statewide Emergency Management Agency (SEMA) in September, the group agreed in their internal meeting to participate in this pilot. Kelly contacted Peter Farmer, head of SEMA IT, who agreed to assist in any way possible. He indicated that the timeline for SEMA to move to Microsoft Exchange has been moved to January 2008. Because of internal consolidation taking place in NC, David Minor indicated that the server is not ready to receive mail but would be ready by January 16, 2008. The server was not ready until February 21, 2008.

#### **Governor's Office— Community and Citizen Services**

The NC team also met with the staff of the Governor's office of Community and Citizen Services. The NC Archives currently receives email from this office every six months via CD-ROM. The group met with the program officer, the Chief Records officer, and the IT support staff to discuss the tool and their participation in the pilot. The staff was very enthusiastic about their participation. They agreed to test the software as well as continue their current sneakernet process should the tool fail. At the same meeting, the NC team received a DVD with 18,973 emails totaling 575,377 KB. As a reference, 10,981 emails were collected in the July-December 2006 time frame from this office.

### **Outreach**

The joint proposal submitted to the Society of North Carolina Archivists by the CERP group and NC Archives and Records was accepted and is on the program for their March conference. In December, the group heard from the Society of American Archivists. The program session that was submitted in September was selected for the August 2008 Annual Convention.

The program committee from NAGARA contacted the group to let them know their submission for the July conference has been accepted.

Megan Durden, a recent graduate of The University of Texas at Austin, joined the NC staff in January 2008. The NC team visited Duke University in February and met with Seth Shaw to talk about efforts in their respective institutions. Seth is very interested in the email parser being developed to access emails from an AOL account. He has the emails but not the software to open them.

### **Other News:**

Kelly Eubank and David Minor spoke with Richard Marciano regarding the tool in development. Richard worked on a much smaller project for the National Archives in 2000 and shared his experiences with us.

KY—Update: Kentucky has placed links to drafts of e-mail training modules on the wiki site. The focus of two of the draft training modules was e-mail management from a records management perspective. These drafts used two separate types of training software. One employed Moodle, an open source course management software that contains course registration, web page presentations, glossaries, quizzes, a structured forum for discussion, and event calendaring. The other draft was created using Softchalk, a commercial course editing tool that facilitates development of web based presentations with hyperlinks, quizzes, and glossaries. Softchalk can produce formatted course materials as a standalone set of web pages or import into course management systems such as Moodle, Blackboard, or Angel. These training tools could also be used to create companion training for H-mail server archival administration, as well as, e-mail system specific training for end users to implement data exchange with the archival hmail server system that the grant is building.

PA—Update: In addition to the TAR approval process moving quite slowly, another unfortunate turn of events for Pennsylvania includes a hold on the e-mail training that was to have been disseminated to all PA Commonwealth supervisor and management employees in December. A conference call was held with OIT/Exchange and all three partners on February 29. Initially, this call was to determine the configuration used in KY since they are also on a MS Exchange mail system. However, the call resulted in little discussion of KY's environment and instead revealed that the OIT/Exchange team member was satisfied with the TAR documentation and that the IMAP connection was appropriate (he was not an original member of the reviewing committee). He suggested a new diagram, which he outlined on a white board, be submitted to the TAR board that he would first endorse. After little movement in the past several months, it appears PA may finally be able to move forward.