Preserving E-Mails with Historical Value

A Case Study

Kurt Brenneman

Government records analysts have one foot in records and information management (RIM) and the other in archives. For example, on the RIM side, they promote best practices and write records retention schedules, and on the archives side, they have responsibility for identifying records with historical value for permanent preservation.

In North Carolina, state records analysts appraise state government archival records – in both physical and electronic formats, including e-mail – based on the records' content. Those with historical value are identified on records retention schedules as having the disposition of transfer to the State Archives of North Carolina (SANC) for accessioning.

Using a Multi-Step E-Mail Retention Process

Because SANC records retention schedules are format-neutral, any records series is likely to include e-mail – but particularly those that include the word "correspondence" in their descriptions. To ensure that state agencies transfer these e-mail messages of historical value

along with their other physical and electronic archival records to SANC, state records analysts insert the phrase "including e-mail" in the series' descriptions and dispositions.

Records analysts also remind North Carolina state employees about e-mail retention in semiannual workshops on e-mail management, instructing staff to flag or isolate e-mails that belong to archival records series.

Like their federal counterparts, North Carolina state employees individually have the burden of determining the retention and final disposition of their e-mail messages based on each message's content and that content's classification on existing records schedules.

State employees are expected to:
1. Categorize e-mail messages according to their content classification on records retention schedules
2. Isolate those with historical value
3. Export e-mails for which archival disposition is indicated into a recommended archival file format
4. Contact a records analyst, who

will transfer these archival e-mails and other electronic and physical records in the records series to SANC using a workflow that protects the records' authenticity

State employees are expected to maintain the integrity of these archival records throughout this multi-step process while also fulfilling the regular demands of their jobs. The success of the process varies, depending on the knowledge, skill, and diligence of the employee performing it.

Looking for a New Approach

With a mandate to preserve the historical record, North Carolina state records analysts recognized the potential of the National Archives and Records Administration's (NARA) Capstone approach as a preservation solution for archival e-mails.

According to NARA's "Guidance on a New Approach to Managing E-mail Records," Capstone captures "records that should be preserved as permanent from the accounts of officials at or near the top of an agency or an organizational subcomponent." Federal agencies can designate Cap-

stone positions at lower ranks when the positions are "likely to create or receive permanent e-mail records."

In other words, archival value is based on e-mail account holders' positions within an organization's hierarchy. All e-mails in e-mail accounts associated with Capstone positions are archival records, no matter their content.

With Capstone, government records analysts apply a uniform retention rule to all e-mails in accounts held by employees in high-level positions. According to William Saffady's January/February 2014 *Information Management* article, "Taking Control of E-mail with Uniform Retention Rules," a uniform rule is easy to understand, requires less decision-making by creators, and is implemented quickly.

Moving Toward Transformation

In 2015, SANC's Digital Services Section proposed to transform the state's management of archival e-mails. SANC, in partnership with archival institutions in other states, received National Historical Publications and Records Commission funding for the Transforming Online Mail with Embedded Semantics (TOMES) project. Among its goals were to create and test new workflows for identifying, transferring, and processing archival e-mails.

SANC would adopt a Capstone-like approach for transferring archival e-mail. Rather than having e-mail account holders categorize, isolate, and export archival e-mails for transfer to SANC, the state's information technology (IT) department staff were to tag the e-mail accounts of state employees whose "Capstone" positions typically have e-mail with significant historical value.

With the approval of the state agencies' chief records officers (CROs), these entire accounts were to be transferred at agreed-upon intervals from the state's e-mail platform to SANC for accessioning and processing.

This workflow was meant to improve the preservation of historically significant e-mail by circumventing the need to rely on the account holders and the limitations of desktop e-mail clients by harvesting e-mails directly from the state's e-mail system based on the account holders' roles.

Launching the TOMES Project

At the TOMES project kickoff meeting, SANC's Digital Services Section and records analysts compared methods for selecting Capstone officials. They considered querying the database of approximately 30,000 records series, extracting archival records series that included correspondence, and tracing them to the positions with custody of those records.

The alternative they chose became a joint effort with state agencies' CROs, who are the ones who approve and sign records retention schedules, represent agencies at SANC's annual meetings, and assist staff with RIM questions. CROs are

A Comparison of Capstone Account Levels

NARA Levels*	North Carolina Levels**
1. Head of agency	1. Head of agency
2. Principal assistants to agency head	2. Principal assistants to head of agency (second tier of management)
3. Deputies to levels one and two	3. Staff assistants to levels one and two
4. Staff assistants to levels one and two	4. Principal management (Chief Operating Officers, Chief Information Officers, Chief Financial Officers, etc.)
5. Principal management (Chief Operating Officers, Chief Information Officers, Chief Financial Officers, etc.)	5. Supervisors and heads of significant divisions, branches, sections, and units
6. Directors or program officers	6. Roles that advise listed positions
7. Regional officials	7. Roles with regulatory approval and rulemaking responsibilities
8. Roles that advise listed positions	8. Additional positions proposed for inclusion
9. Presidential Appointment with Senate Confirmation positions	

* Source: National Archives and Records Administration. "White Paper on the Capstone Approach and Capstone GRS." April 2015.
** Source: State Archives of North Carolina

in high-level positions themselves and are familiar with their agencies' organizational structure.

Choosing an Assessment Form Model

With records analysts' guidance, CROs would identify Capstone positions on a data collection form based upon the NARA-created verification form published in NARA's "White Paper on the Capstone Approach and Capstone GRS."

NARA's verification form captures a list of positions in the nine categories shown in the sidebar "A Comparison of Capstone Account Levels." They begin at the peak of a federal agency's organization and descend eight levels. NARA's form defines the categories, suggests titles in widespread usage, and provides guidance for identifying officials within the categories.

SANC records analysts decided to fashion their own data collection form from NARA's work. They liked that NARA's form was developed for the government sector, reflects the hierarchy of federal agencies, and has good category definitions. North Carolina state agencies, while smaller, are similarly organized and have similar positions and titles.

They also liked that NARA's single verification form fits all federal agencies. According to its "White Paper on the Capstone Approach and Capstone GRS," NARA increased consistency "by ensuring that the same type or level of accounts were identified for permanent retention across the Federal Government."

Adapting NARA's Verification Form

As a first step, SANC records analysts requested organizational charts from each agency's CRO. They noted positions and titles at top levels, studied each agency's records retention schedule for archival records series that included correspondence, and matched records series with the positions.

By doing this, they began to see which levels of agency hierarchies produced archival physical correspon-

dence and the positions that now created archival e-mails. Identifying these levels and positions enabled the records analysts to modify NARA's form to make it applicable for North Carolina agencies. (See the sidebar "A Comparison of Capstone Account Levels.")

To test the waters, records analysts sent SANC's draft assessment form to a handful of CROs, who then convened to discuss it.

SANC's "Assessment of E-mails for Permanent Retention" form asks for specific names, e-mail addresses, and start and end dates of individuals in those identified positions – information needed so the state's IT department could tag their e-mail accounts for the transfer of their e-mails to SANC upon their agency CRO's approval.

The form also asks for names and e-mail addresses of predecessors in Capstone-like positions in hopes of capturing archival e-mails that may have been overlooked.

Though most SANC categories parallel NARA's, SANC records analysts created a distinct category for positions and names of people who were not included in other categories but who also create e-mails concerning historical topics of special interest to SANC. This category is meant to identify Capstone positions not found on organizational charts.

Testing SANC's Assessment Form

To test the waters, records analysts sent SANC's draft assessment form to a handful of CROs, who then convened to discuss it. CROs said they liked the form's clarity and that positions and names could be found in organizational charts and agency

management materials. They felt the assessment form was too lengthy, though, and suggested dividing it into three smaller forms. SANC records analysts reconfigured the form in response to their suggestions:

1. The first form includes the top two levels (i.e., head of agency and principal assistants to the head) and asks for positions of high-level leadership.
2. The second form includes the remaining levels, three through eight, and requests the positions with routine governance of core functions and programs (those that support agencies' goals and missions and serve specific populations and geographic areas).
3. The third form asks for positions not found on organizational charts.

CROs recommended three rounds of data collection as being manageable and were receptive to this new approach to state government records management.

Beginning the Assessment

Each of the three mass e-mailings to the CROs of 19 state agencies (one e-mail for each form) explained the TOMES project and provided a link to a seven-minute SANC video tutorial demonstrating how to complete the three forms. The e-mails emphasized how the project would simplify archival e-mail management, make records creators' jobs easier, and ensure compliance with state public records laws requiring e-mail accounts with archival value to be preserved permanently.

CROs were asked to list position titles, names, e-mail addresses, and start and end dates of employees in the Capstone-like positions. Records analysts provided examples derived from the agencies' organizational charts, from the archival records series in the agencies' records schedules, and from their own knowledge they had gained from schedule writing, consulting, and conducting work-

shops for state agencies. They asked the CROs to review these examples and correct errors in positions and people.

It was important that CROs, rather than records analysts, complete the forms. As senior executives in charge of their agencies' records management programs, CROs know their organizations and often provide insights about their agencies' records to analysts appraising those records. CROs also are the ones to sign off on the transfer of archival records from their agencies.

CROs were asked to return the first assessment form in two weeks. When it was returned, they were sent the second form and asked for its return in three weeks. They received the third form only after completing and returning the first two forms.

Due to the novelty of this approach, these deadlines were considered "soft." Records analysts tracked return dates in a log and sent reminders to CROs who had not responded by the soft deadlines. By checking the tracking log, it was easy for project personnel to quickly assess the TOMES project's progress.

Collecting the Results

CROs from the 19 agencies submitted the first two forms listing 258 positions and a total of 461 e-mail accounts. Records analysts aggregated this information in a single, comma-separated value (CSV) file – making it the official record of positions that created archival e-mails under the TOMES project.

The CSV format is versatile and can be opened and parsed easily, which, according to SANC's "File Format Guidelines for Management and Long-term Retention of Electronic Records," were characteristics SANC and the state's IT department needed to devise an e-mail transfer workflow from the state's e-mail system to SANC.

The quality and comprehensiveness of the responses meant SANC had the names and e-mail addresses for current employees filling the Cap-

stone-like positions, whose accounts needed to be transferred; it had their predecessors' information so it could capture any information that might have been overlooked; and it was well-prepared to tag the accounts of new employees that might move into these positions.

> **… CROs know their organizations and often provide insights about their agencies' records to analysts appraising those records.**

CROs did not return the third form that asked them to list positions that created e-mails concerning specific archival subjects; they did not have enough visibility into these positions. The lack of response to this form confirmed records analysts' belief that only end users can manage and retain e-mail when these decisions must be based on its content.

Listing Lessons Learned

The TOMES project and its Capstone-like method of identifying and transferring 461 state government e-mail accounts with archival value will ensure their preservation. Keys to this project's success were:

1. Identifying the Capstone model on which to pattern the project's approach
2. Adapting the NARA assessment form to fit the needs of the state agencies
3. Collaborating with state agency CROs to take advantage of their organizational knowledge for identifying the positions whose accounts contain e-mails with archival value
4. Developing a tutorial and easy-to-understand instructions with examples for CROs to reference

5. Treating all e-mails of high-ranking positions as a distinct records series, set apart from other related electronic and physical records
6. Eliminating the need for individuals to evaluate each e-mail for its historical significance
7. Leveraging the knowledge and skills of IT staff to develop an automated workflow to transfer the archival e-mail accounts from the state e-mail platform to SANC for preservation

This project proved the contention Saffady made in his article that implementing a uniform rule for e-mail retention can be done "quickly and with a minimum of employee training."

Taking the Next Steps

SANC will arrange periodic transfers of Capstone positions' e-mail accounts – including all metadata – from North Carolina's e-mail system to SANC for preservation.

With the preservation of historically significant e-mail now more assured, SANC will move to test new tools, such as e-discovery software and natural language processing, to make these e-mails more accessible to researchers. **E**

About the Author: Kurt Brenneman is a records analyst with the government records section of the North Carolina Department of Natural and Cultural Resources. His responsibilities include writing records retention schedules and helping North Carolina local governments and state agencies manage public records. Brenneman, who earned a master's degree in library science from the University of North Carolina at Greensboro, can be contacted at *kbrenneman@nc.rr.com*.