Transforming Online Mail with Embedded Semantics (TOMES)
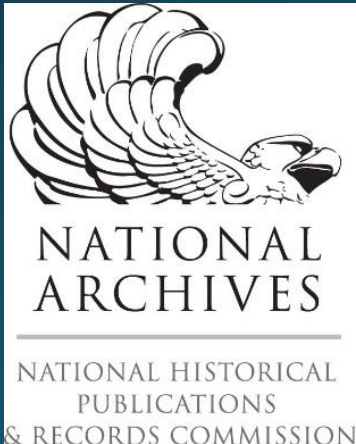
Camille Tyndall Watson

Digital Archivist

State Archives of NC

NATIONAL ARCHIVES

NATIONAL HISTORICAL PUBLICATIONS & RECORDS COMMISSION

NC State Archives of North Carolina

NATURAL AND CULTURAL RESOURCES

Where We've Been...

- NC processes Gov. Jim Hunt emails, 2001-2002
- EMCAP, 2007-2009
- Executive Orders 150 and 18, 2009
- Mimosa Nearpoint, 2009-2014

Where We Are...

- Email retention reduced to 5 years
- State IT has become centralized
- Email has moved to the cloud with Office 365
- Email archive migration

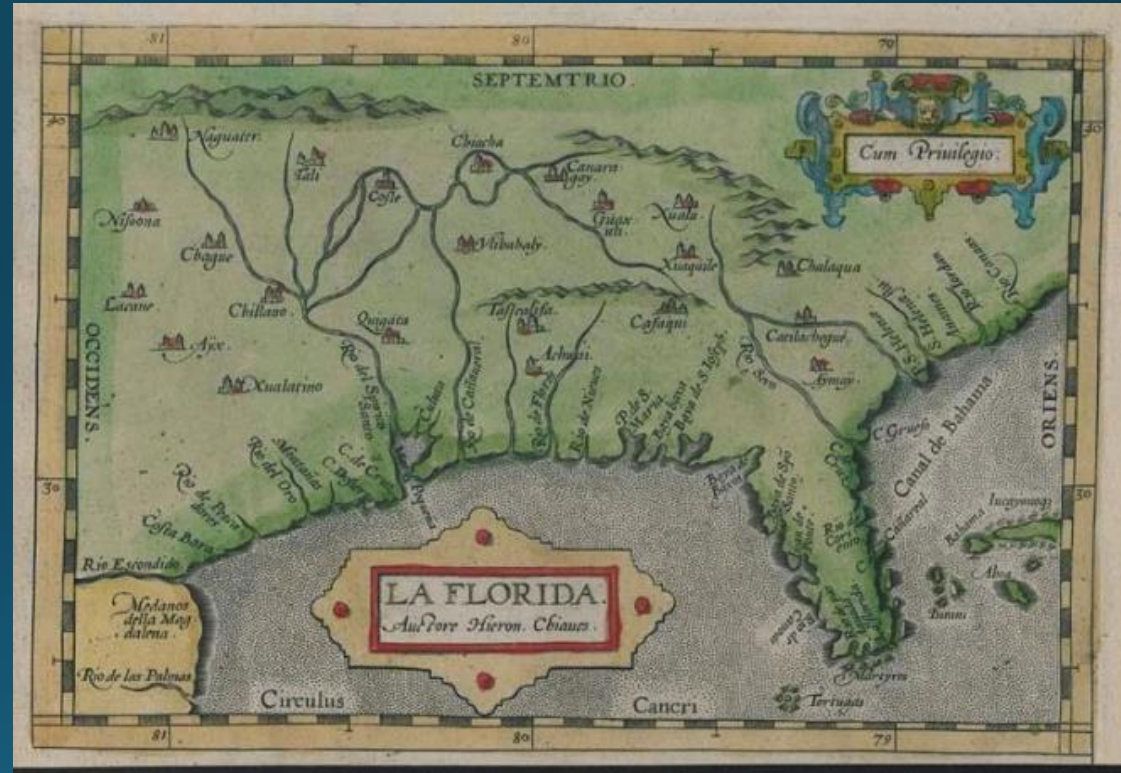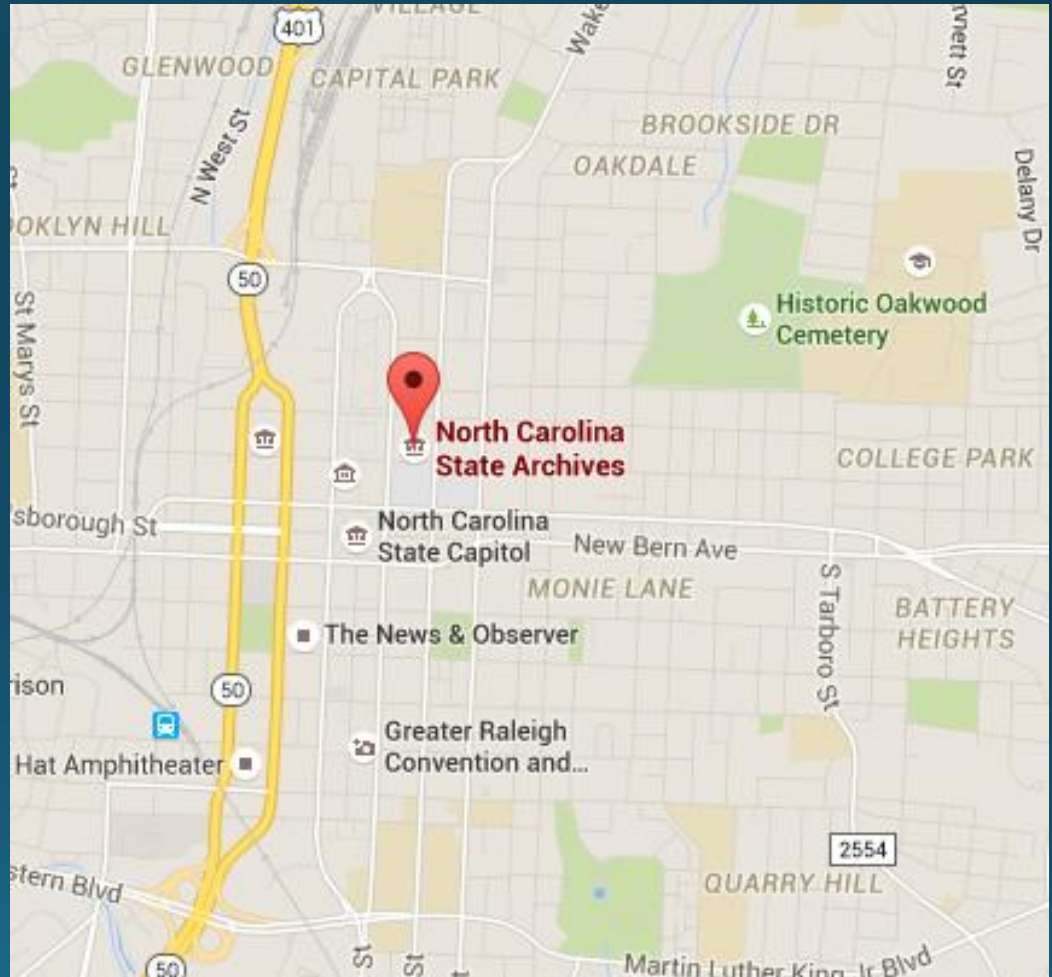Image: Google Maps

Transforming Online Mail with Embedded Semantics (TOMES)

- 3 year grant (2015-2018)
- Partnership between State Archives of NC, Utah State Archives, and Kansas State Historical Society
- Advisory group includes Cal Lee (UNC-Chapel Hill), Chris Prom (University of Illinois Urbana Champaign), and staff from the Library of VA
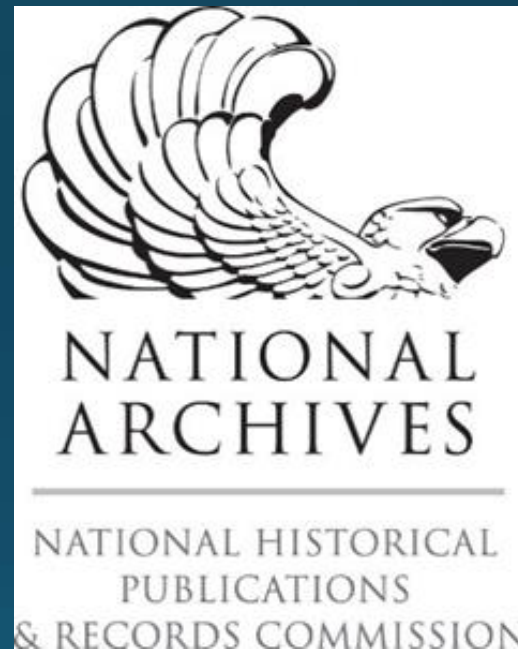
What should we keep?
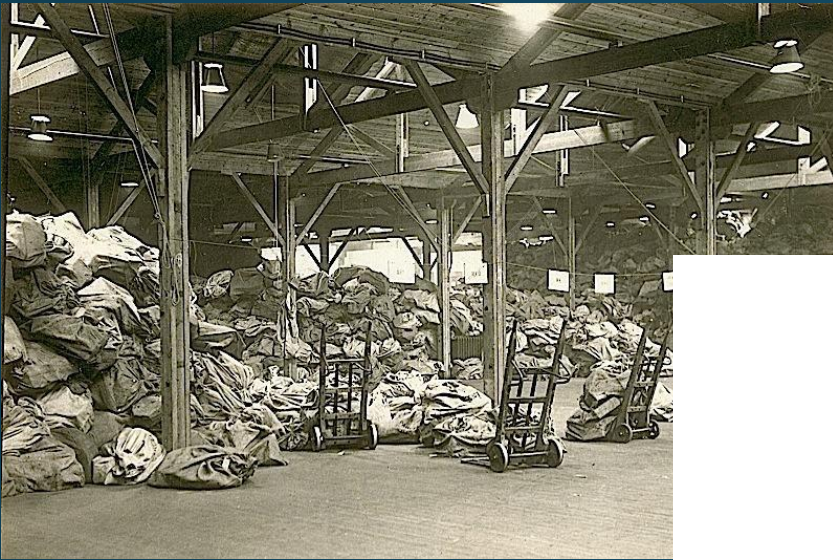
Capstone Approach

Permanent - top decision-maker's e-mail

Temporary — all other staff, 7 year retention

Non-record — 1 year retention

Building Relationships

- Collaboration with Government Records Section, SANC

- Development of a series of forms to understand organizational structures of state agencies

- Outreach to State Agency CROs and CIOs for education and buy in

Collecting Data

State Archives of North Carolina
NATURAL AND CULTURAL RESOURCES

archives.ncdcr.gov

DIVISION OF ARCHIVES AND RECORDS
GOVERNMENT RECORDS SECTION

Agency: _____

Emails from the individuals i...
individuals are high-level, se...

For each individual identified...

1. **The head of the age...**
   president, or equiva...
   although the one pos...

NAME

Predecessors (from...

NAME

State Archives of North Carolina
NATURAL AND CULTURAL RESOURCES

archives.ncdcr.gov   4615 Mail Service C...

ASSESSMENT OF EMAIL...

Agency: _____

Please list below the individuals in the suggested roles and po...
functions and programs of your agency. Core functions or prog...
mission statements and serve significant populations and/or ge...
conduct research supporting core programs, oversee outreach...
wide impact, capture the history of core programs, or are respo...
these email accounts are not already listed under categories 1...

If the emails generated by individuals listed below are already...
the Part 1 form you already completed, including instances wh...
necessary to list those individuals here. Your agency may not l...

For each individual identified below, please list the email accou...

1. **Staff assistants to heads of agencies and their dep...**
   aides. Important work is often carried out by special as...
   behalf of senior officials and/or their email account cont...
   senior officials they support.

NAME	POSITION TITLE/ROLE	B PO N

DIVISION OF ARCHIVES AND RECORDS

State Archives of North Carolina
NATURAL AND CULTURAL RESOURCES

DIVISION OF ARCHIVES AND RECORDS
GOVERNMENT RECORDS SECTION

archives.ncdcr.gov   4615 Mail Service Center, Raleigh NC 27699-4165   919-807-7350

ASSESSMENT OF EMAILS FOR PERMANENT RETENTION
PART 3

Agency: _____

Please identify positions that create official records that document agency policies and
decisions related to any of the programs or subjects listed below. These programs or subjects
are not represented in the accounts of personnel already listed on Parts 1 and 2 but still may
need to be retained permanently. Please note: if emails that document the programs or subjects
listed below are already being captured by the email account of an agency position listed on the
Parts 1 or 2 forms, including instances where the agency executive has been copied on these
emails, it is not necessary to list those positions here. For each individual identified below,
please list the email accounts of all predecessors in that role since January 2011.

Possible email subjects with archival value:

1. Major agency policies
2. Formulation of rules and monitoring standards (e.g., Administrative Code)
3. Events, incidents, and situations that required a prolonged response involving multiple
   agencies and led or had the potential to lead to large-scale loss of life, severe damage to
   lands and property, or major disruption of the state's infrastructure
4. Direction and planning of the core program(s) of your agency
5. Cooperation with external state and/or federal agencies
6. Construction and real property transactions
7. Major public events, such as the State Fair, First Flight Centennial, inaugurations, etc.
8. History of the state of North Carolina
9. Advocacy for minorities, such as Indian tribes
10. Certification, commissioning, etc.
11. Management of assets held in public trust for the people of North Carolina – e.g. state
    parks, historic sites, artifacts, archival materials, etc.
12. Evaluation of rules created by other agencies, where the agency is an established part
    of the rule review process

NAME	POSITION TITLE/ROLE	BEACON POSITION NUMBER	ARCHIVAL SUBJECT MATTER (insert number from list above)	EMAIL ADDRESS	BEGINNING DATE FOR EMAIL COLLECTION

Plugging In

- Department of Information Technology
  - "Tagging" accounts by function
  - Facilitating the transfer of email accounts from cloud storage
- Office of State Human Resources
  - Identifying positions by position number
  - Working with DIT to "tag" accounts

A few bumps in the road…

- Staff changes at partner institutions
- Internal staff changes in DNCR-IT and consolidation with DIT
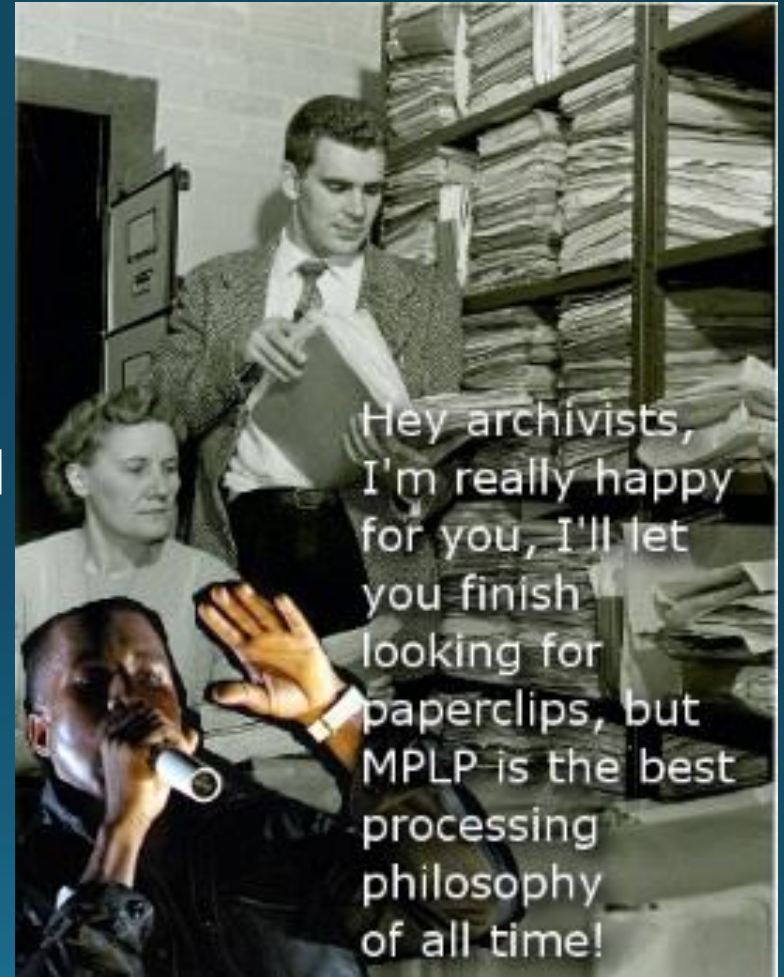- Reorganization of DCR to DNCR
- Changes to division processes
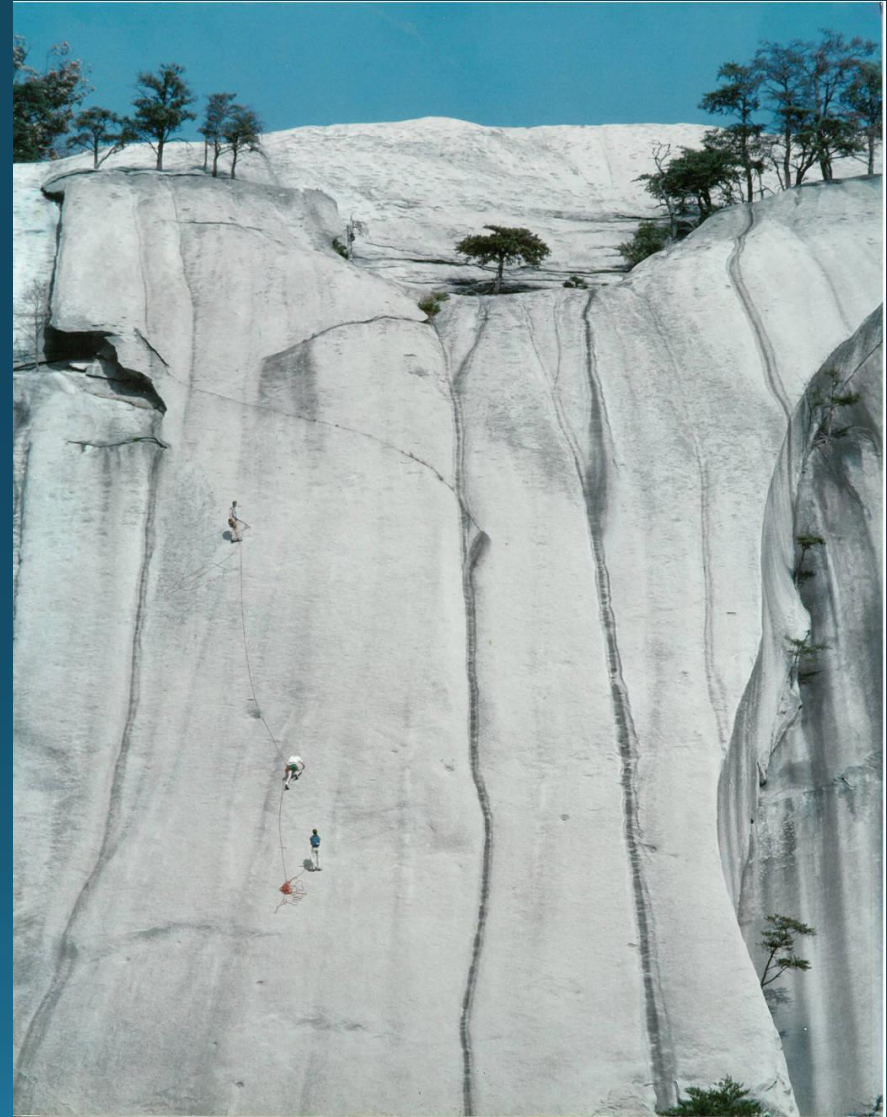
Moving Forward

- Downloading email from cloud using Microsoft eDiscovery tools

- Focus on processing and what we can learn from LVA experience with Gov. Kaine emails

- Programmers began work in Sept 2016

- Converting .pst to .mbox files

- Tool to break .mbox files into more stable sections

- Beginning work on NLP libraries

The Product, After the Process

- Development of State Government specific NLP libraries for use in the processing of email accounts containing public records
- An MPLP approach to the arrangement and description of email
- The ability to identify materials that should be reviewed for PII before release to public

Image: Rebecca Goldman, Derangement and Description, https://derangementanddescription.wordpress.com/2009/10/07/more-meme-less-process/

Stay in touch!

- Git Hub:
  https://github.com/StateArchivesOfNorthCarolina

- Website:

  http://www.ncdcr.gov/tomes
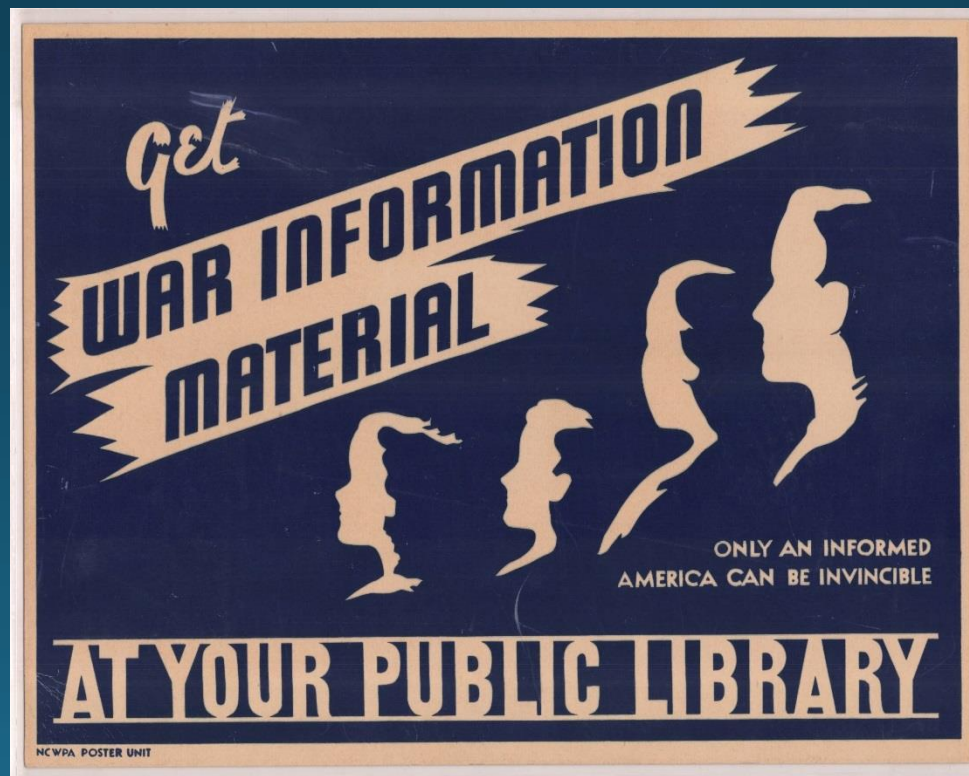
Questions?

Camille Tyndall Watson
Digital Archivist
State Archives of NC
Camille.TyndallWatson@ncdcr.gov
(919) 807 - 7359