

# Using EMCAP (Electronic Mail Capture and Preservation) to Tame the E-Tiger

Glen McAninch, KY

Kelly Eubank, NC

May 29, 2008



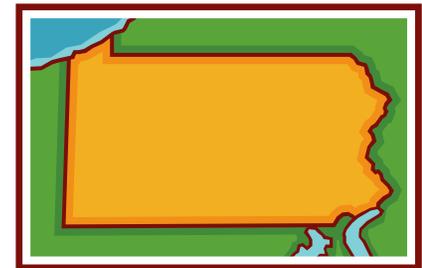
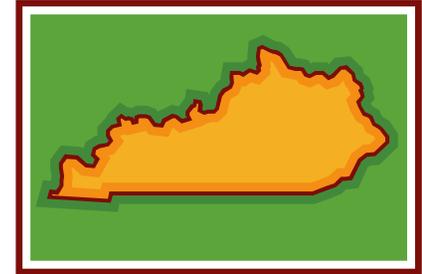
Best Practice Exchange

Helena, Montana



# Email Management Environment

- **Participants—KY, NC, PA,**
- **Diverse systems—centralized versus decentralized**
- **All have retention schedules and training programs**
- **Management of e-mail studied and much debated**



# Transfer Current State of Affairs

- **Manual Collections of Email**
  - Sneakernet, CD-ROM
- **Labor Intensive**
  - No way to stop duplicates, have to run hashing algorithms to determine exact copies
  - Duplicate messages in a variety of accounts
  - Not receiving email from offices
  - Different file formats
- **KY & PA—MS Outlook pst files**



# E-mail Appraisal Activity

- **Kentucky examined 10 years of Libraries and Archives Commissioner e-mail**
- **Kentucky Governor's constituent e-mail conveyed as part of mail tracking databases**



- **North Carolina appraised massive amounts of e-mail from previous administrations**

# Proposed Methodology

- **Build a tool to allow people to “archive” their mail on a systematic basis.**
- **Minimize IT support**
- **Be able to support multiple file formats**



# EMCAP Tool

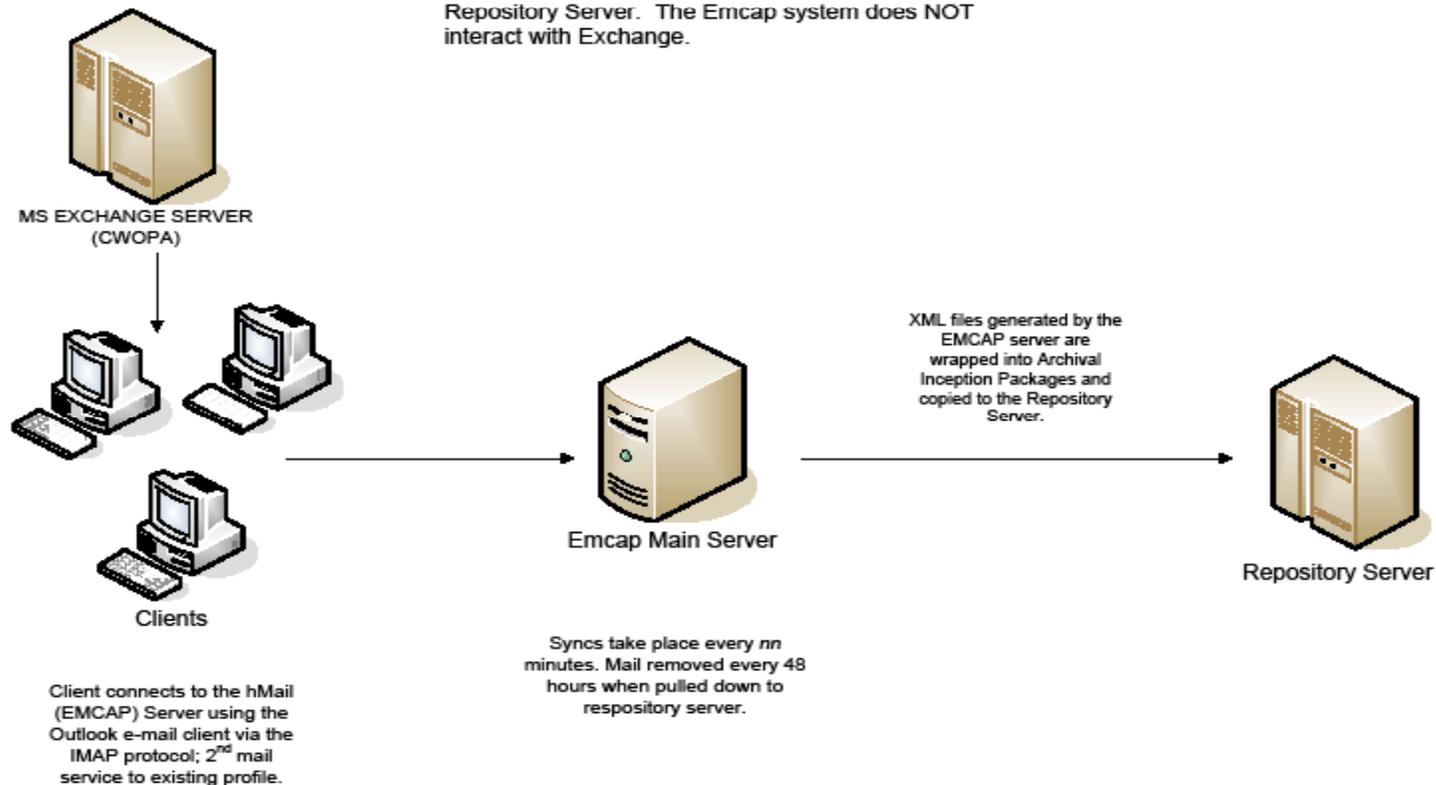
- **Open Source**
- **Client is configured to have “Archives Folder” mapped to server**
- **User can replicate file folder structure**
- **Mimics current behaviors of drag and drop in most e-mail software**
- **Drops email into our collection server**
- **Internet message parser makes an XML copy and also keeps the original**

**Open**

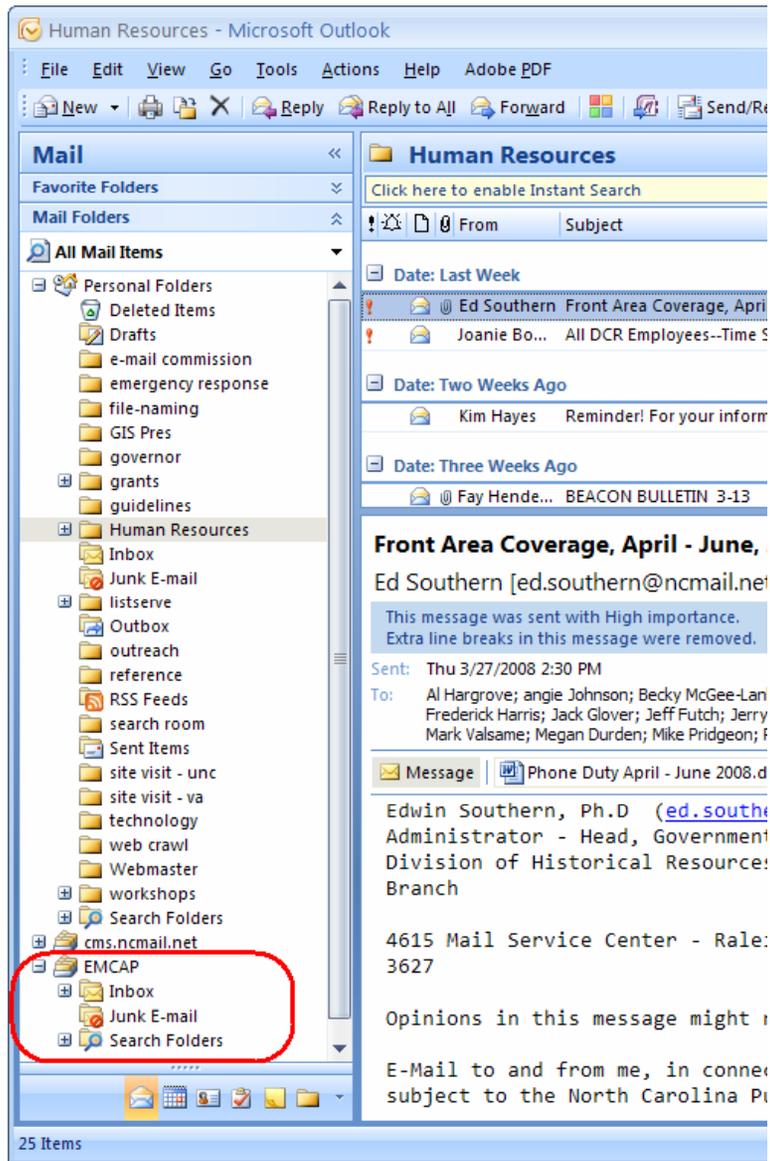
# Architecture of EMCAP

## EMCAP Deployment

In addition to Exchange, clients have IMAP mail service with EMCAP server; Syncs takes place every *nn* minutes. Every 48 hours XML files are generated and both original mail and XML files are pulled down to a Repository Server. The Emcap system does NOT interact with Exchange.



# EMCAP User Client View

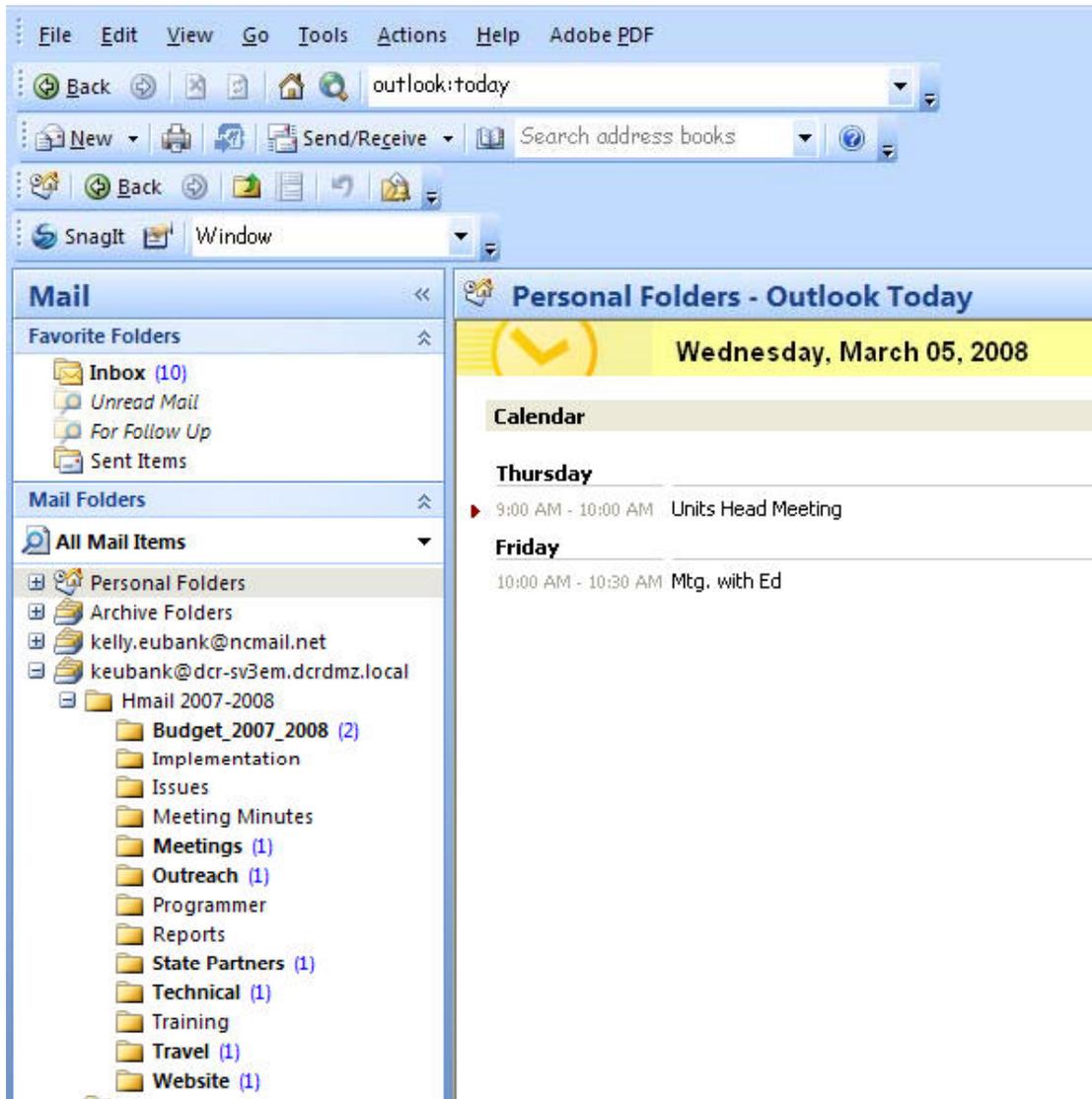


- Can use a variety of software clients

- IMAP connection to the archival server is just another folder in e-mail client you use



# User E-mail moved to EMCAP



- Can be reviewed but not removed by user
- Automatic transfer
- Archivists reviews before creation of XML copies



# Administrative Console

The screenshot shows a Mozilla Firefox browser window titled "Emcap Account Status - Mozilla Firefox". The address bar contains "http://dlaearchive/emcap/". The page content includes the following text and form elements:

**Emcap Account Status**  
Login Page

Kentucky Department for Libraries and Archives  
E-mail Collection and Preservation Server

Please login.

Enter either your archival account address or your regular account address.

Archival E-mail address

Regular E-mail address

Archival account password\*

[Click here if you have forgotten your password.](#)

\* The password used to access your archival account is required here whether you supply your archival e-mail address or your regular e-mail address

The browser's taskbar at the bottom shows the Windows start button, several open applications (Inbox - Microsoft Out..., BPE presentation on t..., Emcap Account Statu...), and the system tray with the time 4:26 PM.

- **Allows for creation and management of accounts**

- **Administrator can reset password**



# Batch Creation of Accounts

Batch Account Creation

File Edit Options

Create Accounts for: GRB and Friends (dcr-sv3em.dcrdmz.local)

Import File:  Load Browse...

	Person Name	Contact Email	Unit Name	Address
*				

Instructions Create Accounts Send Mail Save Report Close

- **Submit list in Excel spreadsheet**
- **Can auto generate e-mail to users**

# Tool will suggest passwords to use

The image shows a Windows operating system interface with three overlapping dialog boxes. The top-most dialog is titled "Account Properties" and is for the domain "GRB and Friends". It contains the following fields: "Account Address" (megan.durden@dcv-sv3em.dcrdmz.local), "Person's Name" (Megan Durden), "Unit Name" (Government Records Branch), "Contact Email" (megan.durden@ncmail.net), "Account Status" (ReadWrite), and checkboxes for "Will Process" and "Active". Below these is a "Notes" text area. The "Change Password" dialog is in the foreground, showing the account name "megan.durden@dcv-sv3em.dcrdmz.local" and empty fields for "Password" and "Confirm Password". It has "Ok", "Cancel", and "Suggest" buttons. A third dialog, "Suggested Password", is open over the "Change Password" dialog, displaying a question mark icon and the text "Would you like to use this password? 4Po\*83". It has "Yes", "No", and "Cancel" buttons. The "Change Password" dialog also has "Help" and "Change Password" buttons.

# User Interface to Set Password

## Welcome To Emcap

Welcome to the North Carolina State Archives  
E-Mail Collecting and Preservation Server.

Jane Doe

Before you can start using your account, you must choose a password.

New Password:

Confirm New Password:

The password must be at least 6 characters in length, and must contain 3 of the following 4 character type: Uppercase letters, Lower case letters, Numbers, and Symbols.

- User has control of password



# EMCAP Account Properties Page

**Account Properties** Domain: GRB and Friends

Account Address:  @dcr-sv3em.dcmdmz.local

Person's Name:

Unit Name:

Contact Email:

Account Status:   Will Process  Active

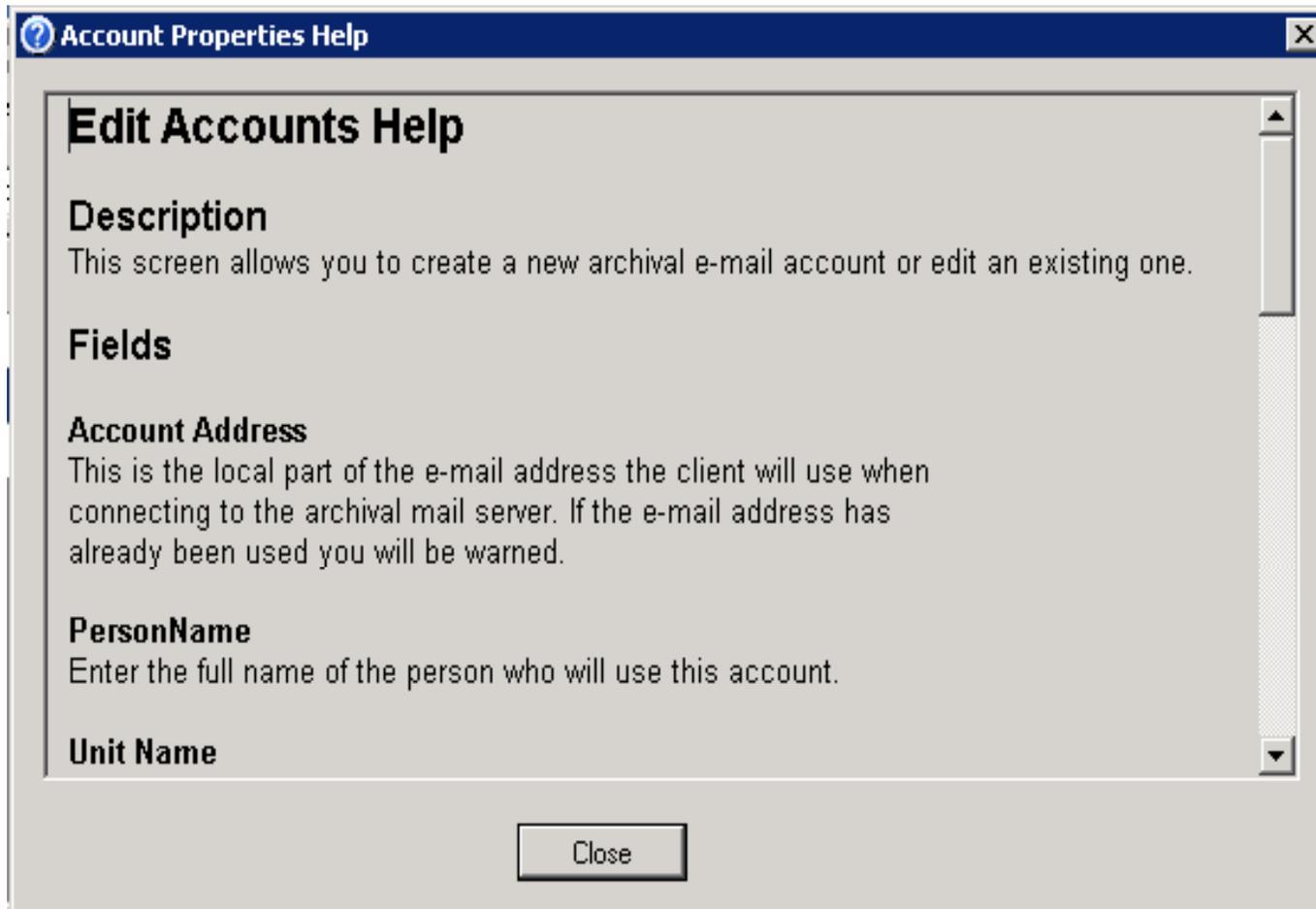
Notes:

**Account Information**

Account Created: <input type="text" value="4/8/2008 7:59 PM"/>	Last processed on: <input type="text" value="Not yet processed."/>
Total Messages Archived: <input type="text" value="0"/>	Messages Archived Last Run: <input type="text" value="n/a"/>
Earliest Message: <input type="text"/>	Online Disk Space: <input type="text" value="Not available."/>
Latest Message: <input type="text"/>	Offline Disk Space: <input type="text" value="0"/>
Store Folder: <input type="text" value="kelly_eubank@dcr-sv3em_dcmdmz_local(1)"/>	

- User and agency info
- Note Field notes
- Storage Totals
- 1<sup>st</sup> & last message

# Help Page for Account Management



- **Help screens explain data entry for all level of user**

# EMCAP XML Creation Tool

- XML schema describes all email in an account
- Develop “common schema” with CERP grant
- Parses the header information in email
- Retain all original bit streams in native format
- When external file save, a message digest is create with a unique identifier
- Preliminary tests show conversion 95% effective



# Example of the XML file created by the Parser

ID and  
format  
info

Text of  
message

From  
and to  
etc.

```
<Name>X-Mailer</Name>
<Value>Unknown (No Version)</Value>
</Header>
- <Header>
  <Name>Message-ID</Name>
  <Value><18a.1823ab0.296c8bde@aol.com></Value>
</Header>
- <Header>
  <Name>X-Filter-Version</Name>
  <Value>1.7 (maildmz)</Value>
</Header>
- <SingleBody>
  - <BodyContent>
    <Content>The Southern Seed Legacy Project based at the University of Georgia has
    completed its web updates. We thank you for allowing us to link to your website. Each
    group has a short paragraph next to their link. Please feel free to comment on these
    paragraphs, if you feel that you would like your group described in another fashion. The
    address for the links page is: http://www.uga.edu/~ebl/links/ If you are interested in
    viewing our home page and seeing what the Southern Seed Legacy Project is all about,
    please check out our main page http://www.uga.edu/~ebl/southernheirloom/ Once
    again thank you. Katie O'Connell University of Georgia Department of
    Anthropology</Content>
  </BodyContent>
</SingleBody>
<Eol>CR</Eol>
</Message>
- <Message>
  <LocalId>0</LocalId>
  <From>"Jennifer Bean" <JBean@oldsalem.org></From>
  <To><ahweb@ncmail.net></To>
  <Subject>Historical Markers</Subject>
  - <Header>
    <Name>Return-Path</Name>
    <Value><JBean@oldsalem.org></Value>
```

# Management of Attachments

- **Links to attachments maintained in native format and converted to Unicode if text**
- **Migration or conversion may be necessary in the future for binary files (pdf, Word, etc.)**
- **System leaves tag in schema to point to the attachment**



# Training Documents

- **User manuals for EMCAP in conjunction with client software**
- **User manual for Administrative Console**
- **E-mail records management training existing in all three states NC, KY and PA**



# Grant Web Site

[www.ah.dcr.state.nc.us/records/emailpreservation/project\\_documentation.htm](http://www.ah.dcr.state.nc.us/records/emailpreservation/project_documentation.htm)

Preservation of Electronic Mail Collaboration Initiative - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ah.dcr.state.nc.us/records/emailpreservation/project\_do Google

## North Carolina State Archives

### Preservation of Electronic Mail Collaboration Initiative

About / Team Members / Project Documentation / Agendas and Minutes / Reports / Online Training / Timeline / Other Email Preservation Sites

#### LINKS

- About
- Team Members
- Project Documentation
- Meeting Agendas and Minutes
- Reports
- Online Training
- Timeline
- Other Email
- Preservation Sites

### Project Documentation

[In Adobe .pdf format]

- Project Summary June 2006**
- Project Narrative June 2006**
- Project Overview September 2006**
- Response to Reviewers September 2006**
- NHPRC Project Report--Preservation of Electronic Mail January 2006**

#### CONTACT INFORMATION

Phone: (919) 807-7355  
Fax: (919) 715-3627  
Email: [kelly.eubank@ncmail.net](mailto:kelly.eubank@ncmail.net)

Mail should be addressed to:

North Carolina State Archives  
4615 Mail Service Center  
Raleigh, N.C. 27699-4615

Physical

**Web site  
contains  
official  
grant  
documents**

**NHPRC**  
DOCUMENTING DEMOCRACY

**Kentucky**  
UNBRIDLED SPIRIT

# EMCAP Use of Web 2.0 Tools

- **Wiki:**



- Storage of draft & official grant documents
- Documentation of technical problems
- Can be used to support the tool in the future

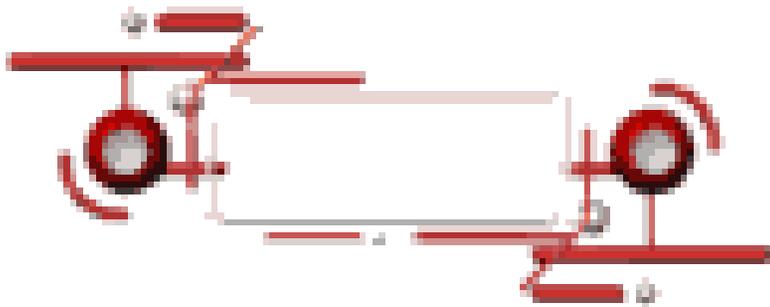


- **Webinar, Course Management and Terminal Services**

- Demonstrations
- Development of training tools

# Next Steps

- **Development of additional file import and export capability – Microsoft .pst format**
- **Test and refine training tools**
- **Implementation by agencies**
- **Refinement of the XML conversion tool**
- **Continue development of tool if funded**



# Points for Discussion

- **Should records creators be the one's to determine what is archival?**
- **How much time can archivists afford to spend on appraisal?**
- **How much technical skill is needed to manage an open source appraisal tool like EMCAP?**



# Usefulness to the Archival Community

- **What kind of e-mail capture is currently being done in your state?**
- **What aspects of the EMCAP tool are most appealing or useful in you in your state?**
- **What other features are needed in an archival tool?**



# Contact Information

---

For more information contact:

Glen McAninch

**Kentucky Department for  
Libraries and Archives**

[glen.mcaninch@ky.gov](mailto:glen.mcaninch@ky.gov)

or

Kelly Eubank

**North Carolina Department of Cultural  
Resources, Office of Archives and History**

[kelly.eubank@ncmail.net](mailto:kelly.eubank@ncmail.net)

