



Smithsonian Institution Archives' Email Preservation Tools

***What they are, how they work and
their operational context***

Riccardo Ferrante

Director of Digital Services

@raferrante

siarchives.si.edu

@SmithsonianArch

Lynda Schmitz Fuhrig

Electronic Records Archivist

@LyndaLSF

Archiving Email Symposium
Library of Congress
Washington, DC – June 2, 2015



Our Starting Point

The screenshot shows a window titled "Elm: IMAP folder hurtta@posti.fmi.fi:RoskaPosti - Shell - Konsole". The menu bar includes "Session", "Edit", "View", "Settings", and "Help". The main area displays the following message list:

Message Number	Date	From	Subject
OM 1	May 19	Travis@mewns12.com	(70) less expensive & No prescription F E E S
O! 2	May 19	rncuka@yahoo.com	(?) URGENT CONTACT gvmk
OM 3	May 19	Thacker	(71) Karaila What meds do u need?
OM 4	May 19	Bernadine F. Ryan	(63) Buy Vicodin at Unbelievable Prices
OM 5	May 19	Erik Weiss	(55) Buy Vicodin Pain Pills Now US Doctors US Phar
OM 6	May 19	Lakeisha Sawyer	(63) unb_eatable software dea-ls
O! 7	May 19	Earnestine Seymour	(?) Top Quality
OM 8	May 19	Angelique Moon	(85) Re:Real Estate Millions evil
OM 9	May 19	Gino Dill	(83) Re:Notice of Default heighten
OM 10	May 19	Rae Kinney	(53) wormy
OM 11	May 19	Aron McCauley	(54) melee
OM 12	May 18	Jenna Hart	(62) stoneware Searching for CHE@P high-quality M!
OM 13	May 19	Inez T. Turner	(60) Free Vicodin Prescriptions.
O! 14	May 19	Bjarne AnderBlanke	(?) Re: Re: Your winning ebay auction
OM 15	May 19	Mohamed Black	(46) limite.d time prices
OM 16	May 18	Comer, Velma	(64) XOXANAX, VALIUM , Cia+lis, VIAGRA and many m
O! 17	May 18	Ulysses Travis	(?) Major Gains-A Story Stock Steal? J k RND_LC_

Below the message list, there is a help section and a command input field:

```
I=pipe, !=shell, ?=help, <n>=set current to n, /=search pattern  
a)lias, C)opy, c)hange folder, d)elete, e)dit, f)orward, g)roup reply, m)ail,  
n)ext, o)ptions, p)rint, q)uit, r)eply, s)ave, t)ag, u)ndelete, or e(x)it
```

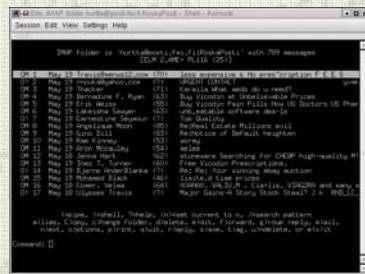
Command: []

Late 1980s ELM (ELectronic Mail)



Our Starting Point

Since then ... an incomplete listing





Our Starting Point

Since then ... an incomplete listing





Our Starting Point

Since then ... an incomplete listing

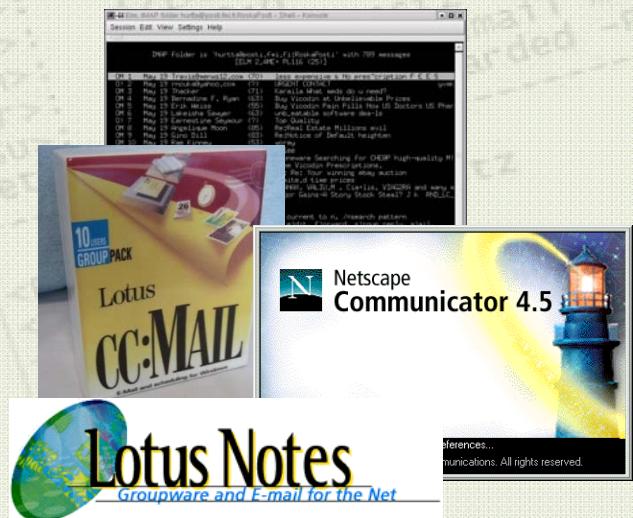


Lotus Notes
Groupware and E-mail for the Net



Our Starting Point

Since then ... an incomplete listing



Novell.

GroupWise™



Our Starting Point

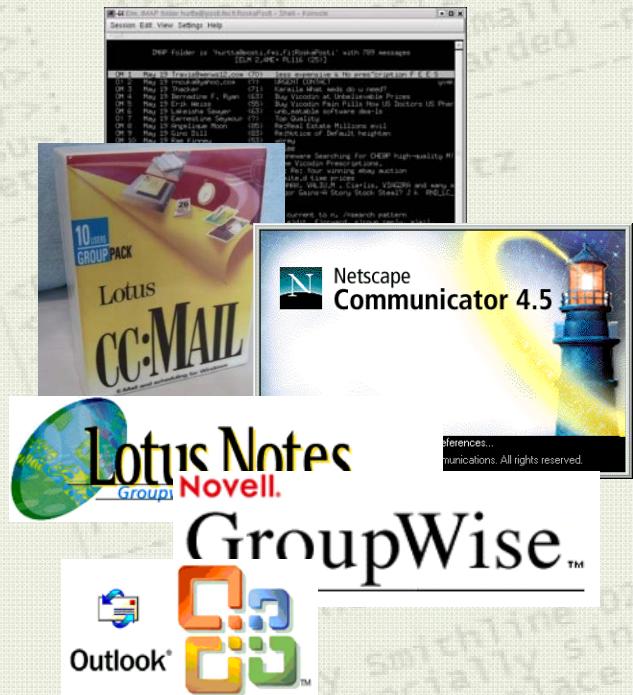
Since then ... an incomplete listing





Our Starting Point

Since then ... an incomplete listing





Our Starting Point



Some of the Smithsonian's email platforms over the past 35 years.

Institutional email management strategies, circa 2004:

- SI Archives-issued guidance part of new employee orientation
- Email typically considered like voicemail
- IT deleted oldest emails on large accounts
- Email migration by special request only when switching from GroupWise to MS Outlook
- Email archives the responsibility of the person
- No established mechanism for identifying and capturing historically valuable emails



Our Starting Point



Some of the Smithsonian's email platforms over the past 35 years.

Late 1980s – the date of some of the oldest email in the Archives collections. At the time, the Smithsonian and many universities were using ELM and PINE.

We do not refuse deposits of email on the basis of format. Therefore we must use an approach that accommodates a variety of formats.

Accessioned email comes from a variety of formats including Elm, Pine, Eudora, ccMail, Lotus Notes, Novell GroupWise and most recently Microsoft Outlook.



The Collaborative Electronic Records Project (CERP)



Some of the Smithsonian's email platforms over the past 35 years.

With these concerns in mind, the Archives teamed with the Rockefeller Archive Center in a three-year project to explore the issue, develop guidance and workflows for acquiring and preserving email.

By the end of the project, CERP had also developed a software tool for preserving email in XML using a schema developed in collaboration with the NC/KY/PA Electronic Mail Initiative project (EMCAP).



Objective: “Durable Access”



- Preservation using non-proprietary, standards based formats
- Asset manageability
- Retention of context
- Retention of complete messages
- Methodology with a low resource requirement



Smithsonian Institution Archives

Our Approach



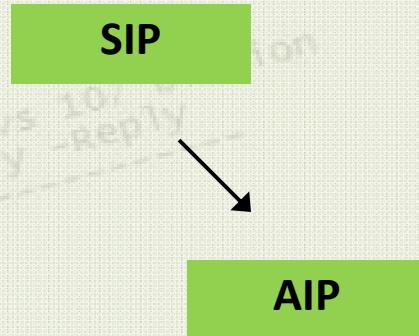
- Since 2008, using the CERP methodology & tools
- Focus on, and actively capturing key organization record creators
- Accounts, not individual emails
- Easier manageability at this level



CERP Model

SIP to AIP

- 1) Convert the collection to .mbox files, if not already in this format.
 - 2) Use the CERP parser to convert .mbox files into an XML preservation file.
 - 3) Assemble all components (metadata, SIP, preservation XML, and log files) into an AIP





CERP Parser

```
<?xml version="1.0" encoding="UTF-8" ?>
<Account xmlns="http://www.archives.ncdr.gov/mail-account"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.archives.ncdr.gov/mail-account/2001/XMLSchema.xsd">
  <GlobalId>204254248
  <Folder>
    <Name>Hash_UTF-8</Name>
    <Message>
      <RelPath>./Steves</RelPath>
      <LocalId>1935189
    <MessageId>
      <![CDATA[<p>051</p>]]>
    </MessageId>
    <MimeType>1.0</MimeType>
    <OrigDate>2001-01-17T12:46:54Z</OrigDate>
    <From>bannerjee@mit.edu</From>
    <To>alevine@mail.csulb.edu</To>
    <Cc>justiga@mail.csulb.edu</Cc>
    <Subject>Newsday</Subject>
    <Header>
      <Name>return-path</Name>
      <Value><![CDATA[<bar>]]</Value>
    </Header>
    <Header>
      <Name>received:</Name>
      <Value>from [128.79.16.58] by mail.csulb.edu with ESMTP id
        f8HGkrs11895; Mon, 17 Sep 2001 12:46:54 -0400</Value>
      <Comments>(bannerjee-mac.csulb.edu [128.79.16.58])(8.11.6/8.11.6)
        (EDT)</Comments>
    </Header>
  </Message>
</Account>
```

E-Mail Account Parsing

All Account directories to be parsed -- and no subdirectories other than properly formed Account directories -- should be located in a directory named "Email_Accounts" which itself must be located in the same directory as the Parser software. Each Account directory must contain all folder subdirectories that you wish included in the parse. Examples might include InBox and/or Sent folders. Any folder may itself contain subdirectories representing sub-folders. Within any folder or subfolder, the file containing email messages to be parsed (one mbox file per folder subdirectory) must be in "mbox" format, with extension ".mbox".

Choose the account directory you wish to parse from the following drop-down list of available candidate accounts that appear to be well-formed.

Once you have chosen the desired target account, press the "Proceed with parsing" button. If the chosen account has already been parsed, you will be asked whether or not you wish to reparse it.

[See help pages for more detail](#)

Choose Account:

Current Parse Status

Parsing initiated 7 January 2015 for C:\EmailParser\Email_Accounts\Carl_backup_messagesave
Begin Parsing: Calendar at 10:18:06 am
34 messages parsed
Done Parsing: Calendar 10:18:08 am
Begin Parsing: DArcInfo Data at 10:18:08 am
1 messages parsed
Done Parsing: DArcInfo Data 10:18:08 am
Begin Parsing: Data Done at 10:18:08 am ... 28 messages parsed

```
<?xml version="1.0" encoding="UTF-8" ?>
<Account xmlns="http://www.archives.ncdr.gov/mail-account"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.archives.ncdr.gov/mail-account/2001/XMLSchema.xsd">
  <GlobalId>204254248
  <Folder>
    <Name>Hash_UTF-8</Name>
    <Message>
      <RelPath>./Hash_UTF-8</RelPath>
      <LocalId>1935189
    <MessageId>
      <![CDATA[<p>051</p>]]>
    </MessageId>
    <MimeType>1.0</MimeType>
    <OrigDate>2001-01-17T12:46:54Z</OrigDate>
    <From>bannerjee@mit.edu</From>
    <To>alevine@mail.csulb.edu</To>
    <Cc>justiga@mail.csulb.edu</Cc>
    <Subject>VTC DNA testir</Subject>
    <Header>
      <Name>return-path</Name>
      <Value><![CDATA[<bar>]]</Value>
    </Header>
    <Header>
      <Name>received:</Name>
      <Value>from [128.79.16.58] by mail.csulb.edu with ESMTP id
        f8HGkrs11895; Mon, 17 Sep 2001 12:46:54 -0400</Value>
      <Comments>(bannerjee-mac.csulb.edu [128.79.16.58])(8.11.6/8.11.6)
        (EDT)</Comments>
    </Header>
  </Message>
</Account>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<Account xmlns="http://www.archives.ncdr.gov/mail-account"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.archives.ncdr.gov/mail-account/2001/XMLSchema.xsd">
  <GlobalId>204254248
  <Folder>
    <Name>Hash_UTF-8</Name>
    <Message>
      <RelPath>./Hash_UTF-8</RelPath>
      <LocalId>1935189
    <MessageId>
      <![CDATA[<p>051</p>]]>
    </MessageId>
    <MimeType>1.0</MimeType>
    <OrigDate>2001-01-17T12:46:54Z</OrigDate>
    <From>bannerjee@mit.edu</From>
    <To>alevine@mail.csulb.edu</To>
    <Cc>justiga@mail.csulb.edu</Cc>
    <Subject>VTC DNA testir</Subject>
    <Header>
      <Name>return-path</Name>
      <Value><![CDATA[<bar>]]</Value>
    </Header>
    <Header>
      <Name>received:</Name>
      <Value>from [128.79.16.58] by mail.csulb.edu with ESMTP id
        f8HGkrs11895; Mon, 17 Sep 2001 12:46:54 -0400</Value>
      <Comments>(bannerjee-mac.csulb.edu [128.79.16.58])(8.11.6/8.11.6)
        (EDT)</Comments>
    </Header>
  </Message>
</Account>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<Account xmlns="http://www.archives.ncdr.gov/mail-account"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.archives.ncdr.gov/mail-account/2001/XMLSchema.xsd">
  <GlobalId>204254248
  <Folder>
    <Name>Hash_UTF-8</Name>
    <Message>
      <RelPath>./Hash_UTF-8</RelPath>
      <LocalId>1935189
    <MessageId>
      <![CDATA[<p>051</p>]]>
    </MessageId>
    <MimeType>1.0</MimeType>
    <OrigDate>2001-01-17T12:46:54Z</OrigDate>
    <From>bannerjee@mit.edu</From>
    <To>alevine@mail.csulb.edu</To>
    <Cc>justiga@mail.csulb.edu</Cc>
    <Subject>VTC DNA testir</Subject>
    <Header>
      <Name>return-path</Name>
      <Value><![CDATA[<bar>]]</Value>
    </Header>
    <Header>
      <Name>received:</Name>
      <Value>from [128.79.16.58] by mail.csulb.edu with ESMTP id
        f8HGkrs11895; Mon, 17 Sep 2001 12:46:54 -0400</Value>
      <Comments>(bannerjee-mac.csulb.edu [128.79.16.58])(8.11.6/8.11.6)
        (EDT)</Comments>
    </Header>
  </Message>
</Account>
```

New Session Configure Toggle Halos Profiler Memory Terminate XHTML 639/309 ms



Preservation and Access

Preservation output
through Email
Account XML
Schema (EAXS)

Human and
machine readable
but not “pretty”

Carl_backup_messagesave(Carl S.aaxml] - <oXygen/> XML Editor

Document Window Help

Saxon-EE

Carl S.aaxml# x

```
46509    </Hash>
46510  </Message>
46511  <Message>
46512    <RelPath>.\Inbox</RelPath>
46513    <LocalId>1190</LocalId>
46514    <MessageId>![CDATA[<A06988C8B0AED4F86EA77B52A3A60BE98A4E08988@SI-MSEV03.US.SINET.SI.EDU>]]</MessageId>
46515    <MimeType>1.0</MimeType>
46516    <OrigDate>2011-09-20T17:49:29-04:00</OrigDate>
46517    <From><![CDATA["SI Email Announcements" <SIAnnounce@si.edu>]]></From>
46518    <Sender><![CDATA["SI Email Announcements" <SIAnnounce@si.edu>]]></Sender>
46519  <To><![CDATA("SI-GEO-FtPierce, FL" <SI-GEO-FtPierce@si.edu>,
46520 "SI-GEO-National Capital Region" <SI-GEO-NCR@si.edu>,
46521 "SI-GEO-NYC All Sites" <SI-GEO-NYC@si.edu>,
46522 "SI-GEO-NZF-SCBI Front Royal, VA" <SI-GEO-NZF-SCBI@si.edu>,
46523 "SI-GEO-Remote Locations" <SI-GEO-RemoteLocations@si.edu>,
46524 "SI-GEO-SAO All Sites" <SI-GEO-SAO@si.edu>,
46525 "SI-GEO-SERC Edgewater, MD" <SI-GEO-SERC@si.edu>,
46526 "SI-GEO-STRI Panama" <SI-GEO-STRI@si.edu>]]></To>
46527  <InReplyTo><![CDATA[<F66362FC0CA446646AC625ADC7FBFE9FF94C3D9AE9@SI-MSEV03.US.SINET.SI.EDU>]]</InReplyTo>
46528  <References><![CDATA[<7A06988C8B0AED4F86EA77B52A3A60BE98A4D11CEC@SI-MSEV03.US.SINET.SI.EDU>
46529 <F66362FC0CA446646AC625ADC7FBFE9FF94C3D9AE9@SI-MSEV03.US.SINET.SI.EDU>]]></References>
46530  <Subject>Event: Health - Seasonal Flu Vaccination Outreach</Subject>
46531  <Header>
46532    <Name>Date</Name>
46533    <Value>Tue, 20 Sep 2011 17:49:29 -0400</Value>
46534  </Header>
46535  <Header>
46536    <Name>x-ms-exchange-organization-authsource</Name>
46537    <Value>SI-MSEHUB-N02.US.SINET.SI.EDU</Value>
46538  </Header>
46539  <Header>
46540    <Name>Thread-Index</Name>
46541    <Value>AQJmQDtxqB2NJ7VK3bra9I4YJgIZxwJNzmqRAqcVRzw=</Value>
46542  </Header>
46543  <Header>
46544    <Name>x-ms-exchange-organization-authas</Name>
46545    <Value>Internal</Value>
46546  </Header>
46547  <Header>
46548    <Name>x-ms-exchange-organization-authmechanism</Name>
46549    <Value>04</Value>
46550  </Header>
46551  <Header>
46552    <Name>acceptlanguage</Name>
46553    <Value>en-US</Value>
46554  </Header>
46555  <Header>
46556    <Name>X-Mailer</Name>
46557    <Value>Microsoft Outlook 14.0</Value>
46558  </Header>
46559  <MultiBody>
46560    <ContentType>multipart/alternative</ContentType>
46561    <BoundaryString>----_NextPart_000_02D0_01CF9D2B.3EA55190</BoundaryString>
46562  <OtherMimeHeader>
46563    <Name>Content-Language</Name>
46564    <Value>en-us</Value>
46565  </OtherMimeHeader>
46566  <Text>\r\n-----\r\nThis is a multi-part message in MHTML format.
```

Text Grid Author

save\Carl S.aaxml Validation successful



Beyond CERP

DArcMail v0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4

Set Account

Account name	Carl_S
Id	11
Account directory	C:\EmailParser\Email_Accounts\Carl_backup_messagesave
Message count	1,226
First message date	2011-09-17
Last message date	2014-05-22
'From' addresses	66
'To' addresses	97
'Cc' addresses	44
'Bcc' addresses	0
External content files	59

Folders

Id	Name	Messa...	Start Date	End Date	Relative Path
359	Calendar	34	2011-10-03	2014-05-14	\Calendar
360	DArcInfo Data	1	2013-05-16	2013-05-16	\DArcInfo Data
361	Data Done	43	2013-01-04	2013-08-15	\Data Done
362	Deleted Items	104	2014-04-04	2014-05-21	\Deleted Items
363	Inbox	616	2011-09-17	2014-05-22	\Inbox
364	Sent Items	424	2011-09-22	2014-05-22	\Sent Items
365	Suggested Co...	4	2012-12-06	2012-12-06	\Suggested Contacts



DArcMail (Digital Archive Mail)

DArcMail v0.1

Quit Load Delete Browse

Account Message Address/Name Results

Date From

Date To

Folder

Subject Line

From Line

To Line

Cc Line

Attachment Name

Body Text

Search body text: text/plain text/html

Reset Search for Messages



DArcMail v0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4 | 5 |

Id	Date	Subject
161950	2012-01-12	RE: Running Post_Init script
162490	2012-01-25	jhove2, droid, etc.
162503	2012-02-21	status
162523	2012-03-22	mysql notes
162045	2012-06-07	Jhove/Droid documentation
162595	2012-09-13	fy 11 numbers for preservation_processing folder on sia-digarch01
162610	2012-11-01	meeting notes
162174	2012-11-08	Working Guidelines for Archiving Websites and Social Media
162224	2013-06-13	FW: Transcription Center in Beta - PLEASE HELP US TEST



DArcMail v0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4 | 5 | 6 | 7 |

Id 162523
Account Carl_S
Folder Sent Items
Global Id <EB83021C560D8E4DAE5EBBC65705487509A4D3B62A@SI-MSEV04.US.SINET.SI.EDU>
Date Thu, 22 Mar 2012 11:10:02 -0400
Subject mysql notes
From "Schaefer, Carl" <SchaeferC@si.edu>
To "Schmitz Fuhrig, Lynda" <SchmitzfuhrigL@si.edu>

Message Content

Action	Storage	Content-Type	Length	Original Name	Par...
view	internal	text/plain	141	-	3
view	internal	text/html	1,831	-	4
download	internal	application/vn...	17,783	mysql_install.docx	5



DArcMail v0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | < | > | X

Here is what we did this morning. We did everything except what is in yellow.

What we installed was version 5.5.21, 64-bit.

Carl



DArcMail v0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12

Dear Colleagues,

The Office of Safety, Health and Environmental Management's Occupational Health Services' seasonal flu vaccination outreach for 2011-2012 has officially begun. On average, Occupational Health Services provides nearly 3,000 shots annually across the Smithsonian. All employees, affiliated staff, and official volunteers are strongly encouraged to get their free flu shot. The complete schedule for this year's outreach effort can be found on the OSHEM website on PRISM:

http://oefo.si.edu/safety_health/ohs/flu_vaccine_fy11.asp

The Centers for Disease Control and Prevention recommend a yearly flu vaccine as the first and most important step in protecting against this serious disease. While there are many different flu viruses, the flu vaccine protects against the three main flu strains that research indicates will cause the most illness during the flu season. The vaccine can protect you from getting sick from these three viruses, or it can make your illness milder if you get a different flu virus.



DArcMail v.0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4 | 5

Id 161886
Account Carl_S
Folder Inbox
Global Id <LISTSERV%201111300908319920.1708@LISTSERV.UCOP.EDU>
Date Wed, 30 Nov 2011 13:08:31 -0400
Subject [JHOVE2TECHNICALTOPICS] MD5 Digest
From "Carl Schaefer" <schaeferc@SI.EDU>
To <JHOVE2-TECHTALK-L@LISTSERV.UCOP.EDU>
Has Replies [Get Replies](#)

Action Storage Content-Type Length Original Name Par...

view internal text/plain 1,187 -

view internal text/plain 6,015 diff.txt

DArcMail v.0.1

Quit Load Delete Browse

Account Message Address/Name Results

1 | 2 | 3 | 4 | 5 | 6 | 7

Id 161891
Account Carl_S
Folder Inbox
Global Id <2DF01F8E2247714A8E141DB29810FF561359C89C@p-irc-exmbx02.AD.UCOP.EDU>
Date Thu, 1 Dec 2011 13:55:04 -0400
Subject RE: [JHOVE2TECHNICALTOPICS] MD5 Digest
From "Stephen Abrams" <Stephen.Abrams@ucop.edu>
To "Schaefer, Carl" <SchaeferC@si.edu>, <JHOVE2-TECHTALK-L@LISTSERV.UCOP.EDU>
In-Reply-To <LISTSERV%201111300908319920.1708@LISTSERV.UCOP.EDU>

Message Content

Action	Storage	Content-Type	Length	Original Name	Par...
view	internal	text/plain	1,657	-	1



DArcMailXML

DArcMailXml v0.1

Convert .mbox Files to XML

Quit

Email Account Information...

Name:

Directory: .

Do only one folder?: yes, limit to:

Split large export into chunks?: messages per XML file

External Content Storage...

Max size for internal storage of attachments:

Create subdirectories for external storage?: yes



Carl_backup_messagesave\Carl Saa.xml - <oXygen/> XML Editor

Document Window Help

Saxon-EE

Carl Saa.xml*

```
</Hash>
</Message>
<Message>
<RelPath>.\Inbox</RelPath>
<LocalId>1190</LocalId>
<MessageId><![CDATA[<AO6988C8BOAED4F86EA77B52A3A60BE98A4E08988@SI-MSEV03.US.SINET.SI.EDU>]]></MessageId>
<MimeType>1.0</MimeType>
<OrigDate>2011-09-20T17:49:29-04:00</OrigDate>
<From><![CDATA["SI Email Announcements" <SIAnnounce@si.edu>]]></From>
<Sender><![CDATA["SI Email Announcements" <SIAnnounce@si.edu>]]></Sender>
<To><![CDATA("SI-GEO-FtPierce, FL" <SI-GEO-FtPierce@si.edu>,
"SI-GEO-National Capital Region" <SI-GEO-NCR@si.edu>,
"SI-GEO-NYC All Sites" <SI-GEO-NYC@si.edu>,
"SI-GEO-NZP-SCBI Front Royal, VA" <SI-GEO-NZP-SCBI@si.edu>,
"SI-GEO-Remote Locations" <SI-GEO-RemoteLocations@si.edu>,
"SI-GEO-SAC All Sites" <SI-GEO-SAC@si.edu>,
"SI-GEO-SERC Edgewater, MD" <SI-GEO-SERC@si.edu>,
"SI-GEO-STRI Panama" <SI-GEO-STRI@si.edu>]]></To>
<InReplyTo><![CDATA[<#F66362FOCA446646AC625ADC7FBFE9FF94C33D9AE9@SI-MSEV03.US.SINET.SI.EDU>]]></InReplyTo>
<References><![CDATA[<#A06988C8B80AED4F86EA77B52A3A60BE98A4D11CEC@SI-MSEV03.US.SINET.SI.EDU>
<#F66362FOCA446646AC625ADC7FBFE9FF94C33D9AE9@SI-MSEV03.US.SINET.SI.EDU>]]></References>
<Subject>Event: Health - Seasonal Flu Vaccination Outreach</Subject>
<Header>
<Name>Date</Name>
<Value>Tue, 20 Sep 2011 17:49:29 -0400</Value>
</Header>
<Header>
<Name>x-ms-exchange-organization-authsource</Name>
<Value>SI-MSEHUB-N02.US.SINET.SI.EDU</Value>
</Header>
<Header>
<Name>Thread-Index</Name>
<Value>AQJmQ0txgB2Nj7VK3bra9I4YJgTZxwJNzjRAqcVRzw=</Value>
</Header>
<Header>
<Name>x-ms-exchange-organization-authas</Name>
<Value>Internal</Value>
</Header>
<Header>
<Name>x-ms-exchange-organization-authmechanism</Name>
<Value>04</Value>
</Header>
<Header>
<Name>acceptlanguage</Name>
<Value>en-US</Value>
</Header>
<Header>
<Name>X-Mailer</Name>
<Value>Microsoft Outlook 14.0</Value>
</Header>
<MultiBody>
<ContentType>multipart/alternative</ContentType>
<BoundaryString>-----_NextPart_000_02D0_01CF9D2B.3EA55190</BoundaryString>
<OtherMimeTypeHeader>
<Name>Content-Language</Name>
<Value>en-us</Value>
</OtherMimeTypeHeader>
<Disclaimers>This is a multi-part message in MTM&lt;br&gt;

```

Text Grid Author

save\Carl Saa.xml Validation successful U+0C



Encoded attachments



Lifecycle coverage from DArcMail Suite

- Processing
- Preservation
 - Everything apart from AIP assembly
- Reference Access (offline)
- Follows CERP methodology



Where this fits in the larger SIA context

Acquisition

- Focused on capturing senior officials, notable persons on departure
- Follows standard electronic records procedures
- Accept unsolicited email records regardless of format
- Work with OCIO to acquire the accounts

Appraisal

- Everything received is preserved
 - Weeding is the responsibility of the account owner
 - Finalized based on subject-sender listing
- Now possible within DArcMail

Arrangement, Description and AIP preparation

- Subject/Sender log is a finding aid supplement
 - Full account structure is embedded in the preserved account's XML.
- Now possible within DArcMail

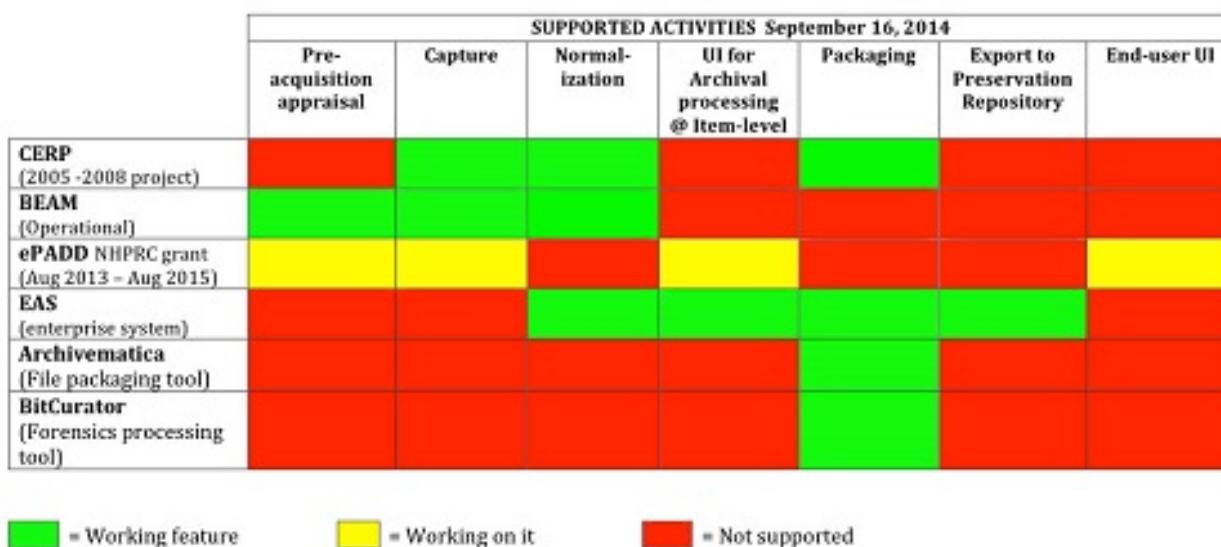
Reference

- Online collection level description
- “Dark archive”
- Restricted for a 15-year minimum



Where CERP fits with other tools

Fulfilling the lifecycle from acquisition to access



CERP Parser handles 3 stages.

DArcMail Suite adds a UI for archival processing and works in Python.

Can be used as part of a composite suite of tools for a full lifecycle solution.

Affordable (open source) and accessible solution for preservation.

