



Smithsonian Institution Archives

# **DArcMail**

---

*Digital Archiving of eMail*

## **Users Guide**

TABLE OF CONTENT

Contents

About DArcMail..... 3

    Issues of scale..... 3

    Open Source..... 3

    Uses for DArcMail ..... 4

    Sample Workflows ..... 5

    Other Notes..... 6

Launching DArcMail ..... 7

    Using DArcMail..... 7

        Launch DArcMail ..... 7

        Logging into the DArcMail database ..... 7

    Navigating the DArcMail Window..... 8

Managing Email Accounts in DArcMail ..... 8

    Loading Data into the DArcMail database ..... 8

Browsing data in the DArcMail database..... 10

    Account ..... 11

    Searching Messages..... 12

    Searching for Addresses or Names ..... 17

    Deleting data from the DArcMail database ..... 19

    Exporting data from the DArcMail database ..... 20

Installing DArcMail..... 21

    Pre-Requisites ..... 21

DArcMailXML ..... 22

    Running DArcMailXml ..... 22

    XML file names..... 23

    MessageSummary.csv output ..... 23

CmdDArcMailXml..... 24

    Running CmDArcMailXml..... 24

## About DArcMail

The **DArcMail** application builds upon the functional core of the Collaborative Electronic Records Preservation (CERP)<sup>1</sup> Parser. DArcMail delivers a broader range of features and functions and does so on the popular Python platform. In addition to preservation functionality, DArcMail supports appraisal, processing, and providing access to email accounts or large groups of related email messages. DArcMail's design shrinks processing time and enables it to handle large email accounts on the three most popular operating systems: Windows, Mac OS, and Linux systems.

Organizations that are only interested in preserving email accounts to Email Account<sup>2</sup> compliant XML may prefer to use **DArcMailXML** (graphical UI) or **CmdDArcMailXML** (command line UI).

### *Issues of scale*

In digital preservation circles today, the effective capacity of a tool or a workflow is sometimes described as handling “large amounts” of email. Without actual numbers, a capacity to handle “large amounts” of email is meaningless and does not help the digital preservation practitioner or administrator select a tool suited for their own context.

For clarity's sake, DArcMail is designed to scale, handling anything from a single email message to an account containing over a quarter of a million messages with ease. The Smithsonian Institution Archives regularly accessions the email correspondence of key Smithsonian staff, acquiring their whole email account at the time when such a staffperson leaves the Institution. More often than not, Smithsonian museum directors, senior curators and senior scientists may have more than 100,000 emails in their account when acquired.

### *Open Source*

**DArcMail** and **DArcMailXML** are available as open-source software under the MIT license (<http://opensource.org/licenses/MIT>). Its two platform components, Python and MySQL/SQLite, are also open source, continuing in the spirit of the CERP and the Smithsonian Institution Archives commitment to sharing digital

---

<sup>1</sup> <https://siarchives.si.edu/oldsite/cerp/index.htm>

<sup>2</sup> Email Account XML Schema

[http://www.history.ncdcr.gov/SHRAB/ar/emailpreservation/mail-account/mail-account\\_docs.html](http://www.history.ncdcr.gov/SHRAB/ar/emailpreservation/mail-account/mail-account_docs.html)

preservation resources. The DArcMail application utilizes Python and MySQL for the full set of functionality. Both Python and MySQL are well established open source tools available on Windows, Macintosh and Linux platforms.

## Uses for DArcMail

DArcMail is intended primarily for the processing, preservation migration and subsequent access of email accounts or groups of related emails:

- Appraisal

Searching through an email account using DArcMail helps an archivist develop a more detailed knowledge of the account owner's use of email, their correspondents, focal topics, etc. Simple and complex searching is available along with filtering and thread tracking features. Individual messages can be identified for exclusion before finalizing the set of emails to be accessioned. A log of senders and subjects can also be generated should the archivist find this helpful in making their determination.

- Processing

Like its usefulness for appraisal, DArcMail's processing capabilities lie in its ability to identify groups of messages based on a wide variety of fielded criteria. Correspondents, subject and body text content, topics, time frames, email threads, attachments – all of these criteria can be used to examine, filter and analyze the full email account or subsets of its messages. The knowledge gained will facilitate the creation of a detailed finding aid or narrowing the scope of the messages to be formally acquired by the organization.

- Preservation

DArcMail's preservation function transforms the original MBOX-formatted email account into XML as the preservation master format of the accessioned email account. A single XML file can be generated for the entire account, organized to also capture any hierarchical organization present in the original account.

The XML format is compliant with the Email Account XML Schema<sup>3</sup> and can easily be validated by most XML editors. The EMA-compliant file captures every aspect of the email account: 1) internal folder structure and relationships; 2) email message content, associated transmission headers; and 3) messages and

---

<sup>3</sup> ibid

their attachments<sup>4</sup>. In addition, DArcMail generates a Message Summary log which contains the fields for Date, To, From, Message ID, Subject, and SHA1 Hash<sup>5</sup>.

- Access

DArcMail can be used to provide access to email collections in a reading room setting or for reference staff. The relevant functionality appropriate to this use is the same searching and filtering capabilities described in the previous Appraisal and Processing sub sections.

### *Sample Workflows*

#### Acquisition to Preservation

1. Email account is chosen for appraisal
2. The original account file is scanned for viruses
3. If the email account file format is not MBOX, the format is normalized to MBOX.
4. The MBOX-formatted account is imported to DArcMail.
5. The archivist reviews the account's messages in DArcMail and exports the email messages that will be accessioned as a set into an MBOX formatted file.
6. Archivist uses the DArcMailXML component to perform the preservation migration from the accessioned MBOX file to the preservation XML file along with the MessageSummary log and a directory of any associated message attachments

#### Processing / Access

1. Email account's MBOX file is imported into DArcMail.
2. Archivist searches through and reviews messages in DArcMail.
3. Note: If the archivist wishes to provide a subset of messages to a researcher, they will create a filtered set of messages using DArcMail as a derivative MBOX file.

---

<sup>4</sup> DArcMail can be configured to embed attachments directly in the XML generated or to move attachments to a parallel directory structure, leaving pointers in the XML to reference the moved attachments.

<sup>5</sup> A message-specific fixity value using the SHA1 algorithm. <https://en.wikipedia.org/wiki/SHA-1>

## Other Notes

**Message content.** The content of any given message may be small or large. With multiple attachments, the total content size for a single message can exceed 10 megabytes. Storing all message content for a large account that may contain more 100,000 messages in the database may be unwise. The main reason for this is not the number of messages, but the potential size of their parts. For this reason, DArcMail allows you to divide the storage for message content between the database and external storage. Here are the rules:

All header information for all messages is stored directly in the database. The content of any given message may be stored in the database or externally.

The MBOX format divides the content of a message into "parts." In the DArcMail database, each message part is stored separately; the maximum size for database storage of a single message part is 60,000 bytes.

When loading an email account (or loading one email folder of an account), you may specify the upper size limit for database storage of a message part. If you specify 0 for this upper limit, then all parts of all messages will be stored externally. If you specify 60,000 for this upper limit, all parts that are 60,000 or smaller will be stored in the database.

**Account directory** is the location of all externally stored content such as message parts that exceed the size specified when the account is loaded into DArcMail.

.raw files are produced by the DArcMail loading process. These are files for externally stored message parts. This occurs when you load up an account and set a max internal storage per part (up to 60K). Any part exceeding that size gets stored externally in a RAW file

.xml files for email attachments are produced by the DArcMailXml conversion process. For example, a Microsoft PowerPoint (ppt) file that was encoded in base 64 to start would be xml encoded, remaining in base 64.

The account directory may contain multiple subdirectories; for example, there may be a separate subdirectory for each email folder in the account. But while subdirectories are recommended for larger accounts, this is not necessary.

**Folder directory.** Every email account will have one or more folders, with one MBOX file for each email folder. For a given email folder, the directory in which the MBOX file for that folder is stored is the "folder directory" for that folder. If the MBOX file is located directly under the account directory, then the account directory and the folder directory

are the same. If you want, you can place multiple MBOX files in the same directory, although for larger email accounts we recommend having each email folder in a separate subdirectory of the account directory.

## Launching DArcMail

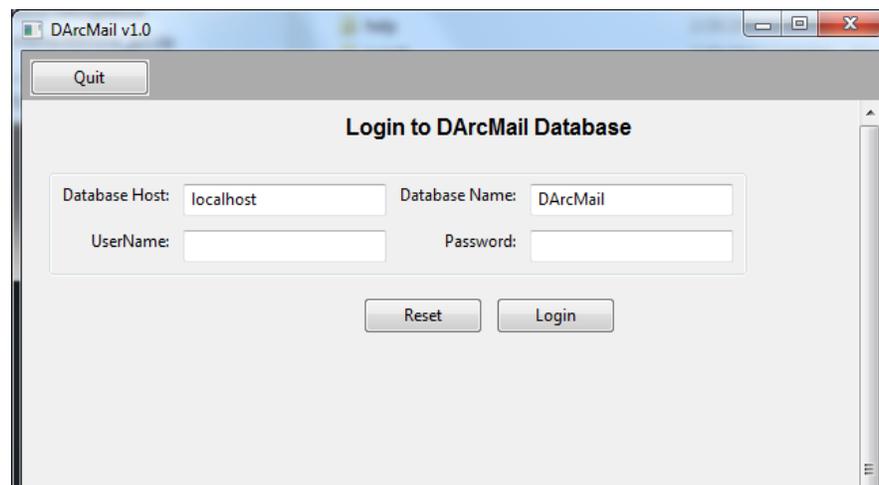
### Using DArcMail

#### *Launch DArcMail*

The main DArcMail program is the file DArcMail.py. Create a shortcut for this file and place it on your desktop. If the .py extension has been associated with the Python interpreter (on Windows) or with the Terminal application (on Mac OS X), then you can double-click the file to start the program. If this has not been done, do so following the operating system specific instructions available from Microsoft or Apple. Otherwise, you can start the program from a Windows command prompt window or from a Mac OS X Terminal window.

#### *Logging into the DArcMail database*

The login screen is the first screen that appears when you start the DArcMail program.



If you are using DArcMail with a SQLite<sup>6</sup> data file, leave the database field blank. No username or password is required.

---

<sup>6</sup> <https://www.sqlite.org/about.html>

If you are using DArcMail with a MySQL database backend, the Database Name will be the name of the database in MySQL. Your "Username" and "Password" are your username and password for that MySQL database.

If you are running the DArcMail.py program from a computer on your network, the Host name is the network name of the computer running MySQL. If you are running MySQL on the same machine as DArcMail, use "localhost."

## Navigating the DArcMail Window

After you have successfully logged into the DArcMail database, then the DArcMail.py program will show buttons at the top of the window for exiting, loading, deleting, exporting and browsing data.

## Managing Email Accounts in DArcMail

### Loading Data into the DArcMail database

The "Load DArcMail Data" screen asks for the following information:

The screenshot shows the DArcMail v1.0 application window. At the top, there is a menu bar with buttons for 'Quit', 'Load', 'Delete', 'Export', and 'Browse'. The main window title is 'Load DArcMail Data'. Below this, there is a section titled 'Email Account Information...' containing several input fields: a 'Name:' text box, a 'Directory:' text box with a 'Browse' button, and a checkbox labeled 'Do only one folder?:' with 'yes, limit to:' and an empty text box. Below this is another section titled 'External Content Storage...' which includes a dropdown menu for 'Max size for internal storage:' set to '30000' and a checkbox for 'Create subdirectories for external storage?:' which is checked and labeled 'yes'. At the bottom of the window are two buttons: 'Reset' and 'Load Data'.

**Name.** This is the DArcMail name of the email account. This can be the email address that is associated with the email account (e.g., "john.smith@jamestowncolony.com") or any other format (e.g., "John Smith", "Smith, J; Jamestown", etc.). If the email account name does not already exist in the DArcMail database, it will be created in the loading process.

**Directory.** This is the directory where the MBOX file is located. This must already exist and must contain (directly or indirectly) one or more MBOX-format files with the .mbox file extension.

**Folder.** If no folder is specified, then the program will load message data from all the MBOX files that are contained below the account directory. If the "yes" box is checked and a folder "AbcXyz" is specified, then the program will load message data only from the MBOX file named "AbcXyz.mbox".

Note: A folder name is taken directly from the name of the MBOX file. Each folder name in an account must be unique. In the following example, attempting to load account JohnSmith, with account directory "G:\EmailData\JohnSmith", would fail since the folder name "Inbox" is not unique within the account directory:

```
G:\EmailData\JohnSmith\archive\Inbox.mbox
```

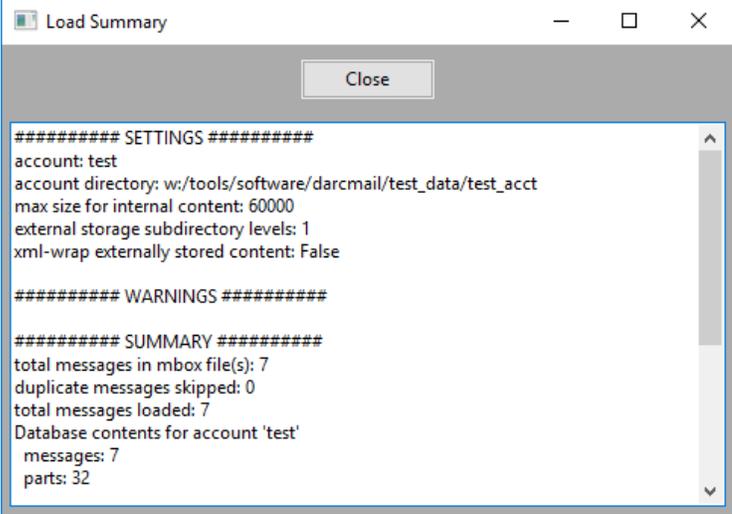
```
G:\EmailData\JohnSmith\current\Inbox.mbox
```

Note: If the specified folder has already been loaded, you must first delete it before reloading.

**Max size for internal storage.** This sets the upper limit, in bytes, of the size of a message part if the part is to be stored internally in the database. This particularly applies to message attachments. See message content storage for further information.

**Subdirectories for external storage.** If "yes" is checked, then message content that is stored externally will be distributed among subdirectories of the applicable folder directory. This feature replicates the organizational structure of the account in the external directory. Use it to prevent the creation of directories with thousands of files. See message content storage for further information.

Once loaded, a summary will be displayed, such as:



```

Load Summary
Close

##### SETTINGS #####
account: test
account directory: w:/tools/software/darcmail/test_data/test_acct
max size for internal content: 60000
external storage subdirectory levels: 1
xml-wrap externally stored content: False

##### WARNINGS #####

##### SUMMARY #####
total messages in mbox file(s): 7
duplicate messages skipped: 0
total messages loaded: 7
Database contents for account 'test'
  messages: 7
  parts: 32

```

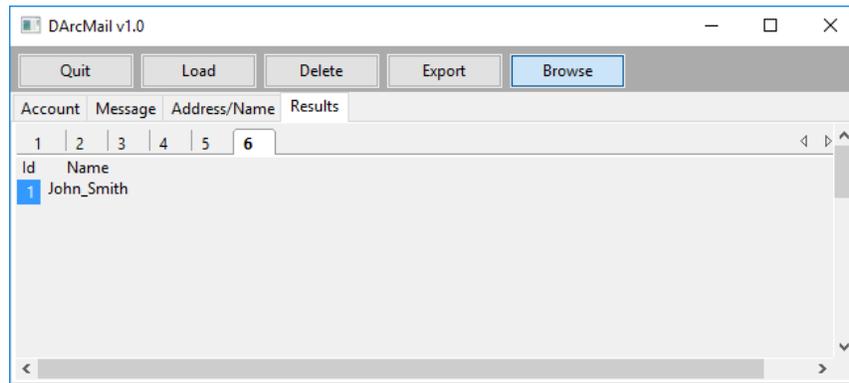
## Browsing data in the DArcMail database

You can search for accounts, for messages, and for email addresses or email names. The result of a search is an account list, a message list, or an email address/email name list. Selecting one row of a list will result in an information page for the individual account, message, or email address. Lists and information pages appear under the Results tab. A list or information page remains under the Results tab until you delete it.

When you are searching for a message or an email address, your search is restricted to a single current account. You set the current account by pressing the "Set Account" button on the account's information page.

There is a 1,000 results limit on the number of items that can be displayed in a list; if an account list, email address list, or message list exceeds 1000, only the first 1000 accounts, addresses, or messages will be shown. If the result list is longer than 30 items, additional tabs will be created for each additional group of 30. This is done in order to preserve low response times.

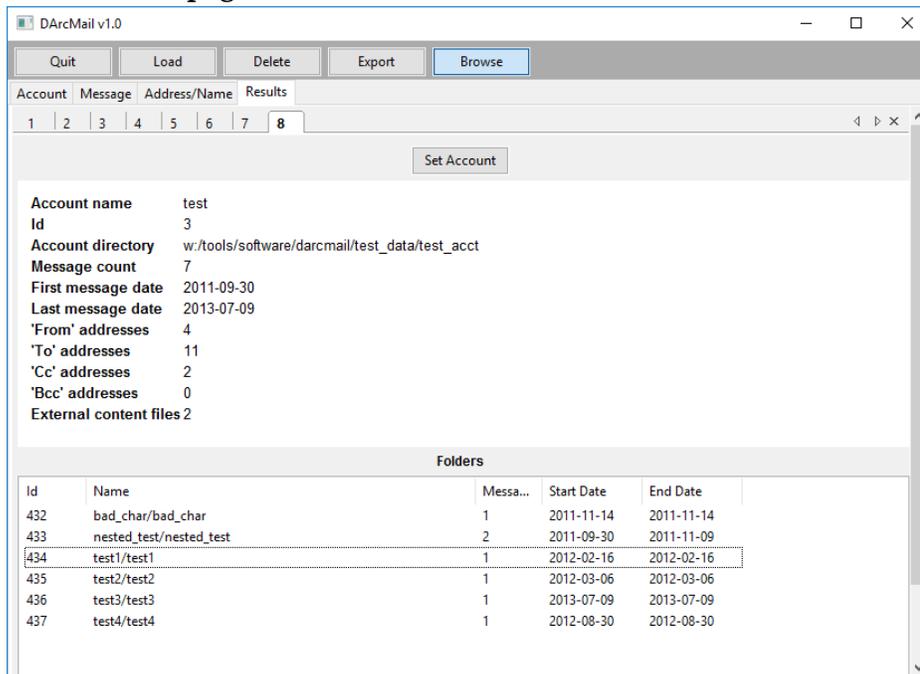
You will see four tabs:



## Account

You will first identify the account you want to work with. You can search on a partial or complete account name. The search string you supply will be automatically wild-carded at the beginning and end. Searches are not case sensitive. If you do not supply any search string, the search will match all accounts.

## Account info page



The account info page has three components:

The "Set Account" button, *Account attributes*, and a *List of Folders*. Pressing the *Set Account* button will set the current account for any subsequent message or address search. Click on the Message to open a Message Search window. To open an Address/Name Search window, click on Address/Name.

## Searching Messages

You can search for messages using any combination of the search boxes in the search form. If you supply values for two or more of the search boxes, then the search must match all of the values you supply. This behaves like a Boolean "AND" search query.

All text fields are automatically wild-carded at the beginning and end.

The "Date From" is wild-carded as follows: A year only entry "2014" is interpreted as "2014-01-01." A year-month entry "2014-03" is interpreted as "2014-03-01".

The "Date To" field is wild-carded as follows: A year only entry "2014" is interpreted as "2014-12-31." A year-month entry "2014-03" is interpreted as "2014-03-31".

"From", "To", and "Cc" fields will search both email addresses and email names (see Address for treatment of addresses and names).

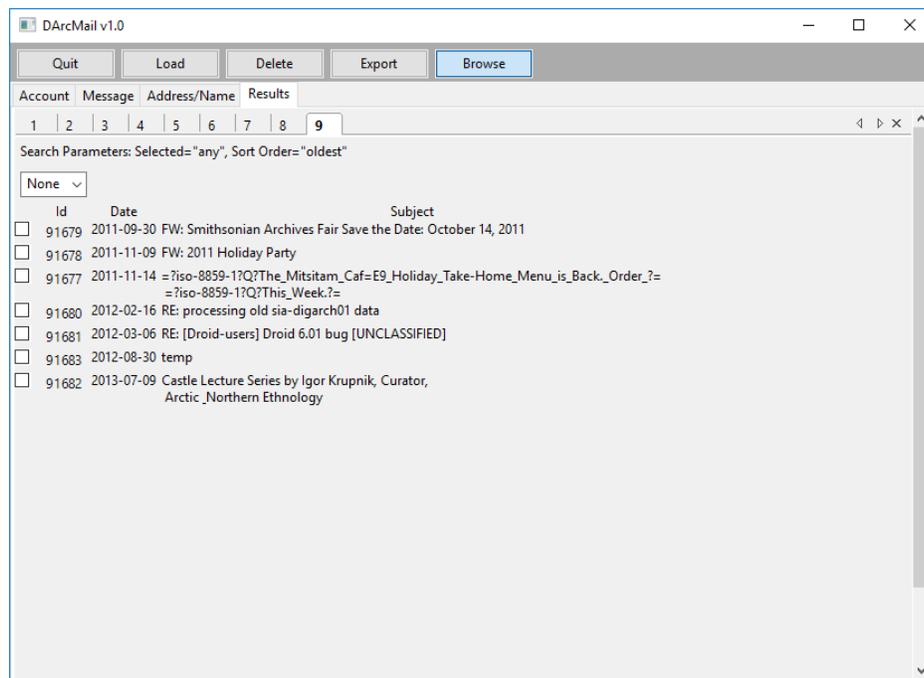
Text searches are not case sensitive.

You can search on attachment names even if the attachment itself is stored externally. As a text field that is automatically wild-carded, if you supply the string ".jpg" in the "Attachment Name" field, the search will find all messages with at least one part that is an attachment having ".jpg" in the attachment name.

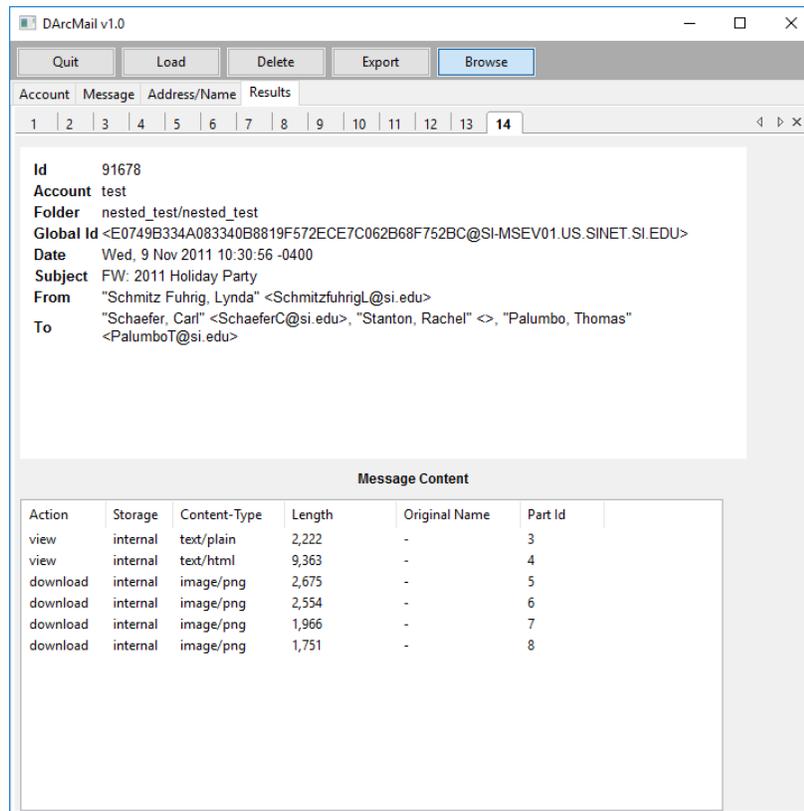
If you supply a search string for the "Body Text" field, only messages with Content-Type text/plain and text/html will be searched, and only content that is stored internally will be searched.

Note: Message content is stored (and therefore searched) in its original transfer encoding. Content-Type 'text/html' is usually encoded as 'quoted-printable'; this may affect your search results. In addition, if you specify a search of 'text/html' content, your search may match html tags which are not in fact visible in the rendered html. A message part with Content-Type 'text/html' is usually accompanied by a part containing the same information with Content-Type 'text/plain.' For all these reasons, you will probably prefer to restrict your search to 'text/plain.'

If you do not supply any search string, the search will match all messages.



## Message information page



A message information page has two parts:

The top part of the page shows some attributes of the message, including information from these MBOX header lines "Message-Id:", "Date:", "Subject:", "From:", "To:", and if present, "Cc:", "Bcc:", and "In-Reply-To:". If the message has an "In-Reply-To:" header and if the replied-to message is in the DArcMail database, then the message id of the replied-to message is hyperlinked to its message info page. If the message itself has replies, then the text "Get Replies" (or "Get Reply") is hyperlinked to a list of the replying messages.

### Message Content List

An MBOX message may have just one body part or it may have multiple body parts (with a Content-Type of multipart/alternative or multipart/mixed or multipart/related or multipart/report, etc.) If a message is multipart, then its first part is simply the multipart container whose contents are the other parts. A

multipart message part (a "container" part) does not appear in the message content list at the bottom of a DArcMail message info page.

The message content list shows these fields for each non-empty, non-multipart message part:

- Action. This is 'view', 'download', or 'show path'.
  - o If the Content-Type of the message part is 'text/html' or 'text/plain' AND either
    - (a) the Storage is 'internal' OR
    - (b) Storage is 'external' and the external file is visible from the computer on which the DArcMail application is running, and the external content has NOT been wrapped in XML,
 then the action is 'view', meaning that you can view the content directly in the DArcMail application.
  - If the Content-Type is not 'text/plain' or 'text/html' AND either
    - (a) the Storage is 'internal' OR
    - (b) Storage is 'external' and the external file is visible from the computer on which the DArcMail application is running,
 then the action is 'download'.
  - If the Storage is 'external' and the external file is NOT visible from the computer on which the DArcMail application is running, then the action is 'show path', meaning that DArcMail will show the full path of the external file in a separate window.

If a message part has a Content-Transfer-Encoding of 'base64' or 'quoted-printable', DArcMail will decode the content to its original representation before viewing or downloading. However, if the Storage is external and if the external content has been wrapped in XML, then the content will not be decoded.

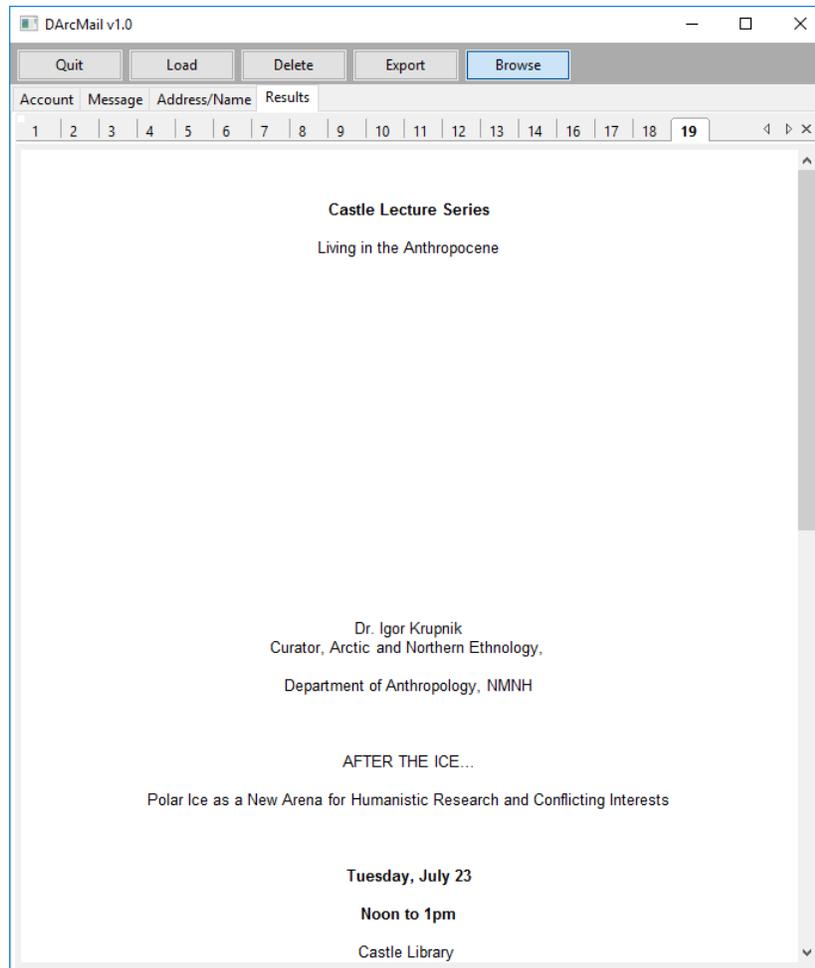
**Content-Type.** This is the MIME type of the message part, as given in the Content-Type header of the message part. If a message part has Content-Type 'text/html' the preceding or following message part may be a 'text/plain' rendering of the same information.

**Length.** This is the length of the content, in bytes.

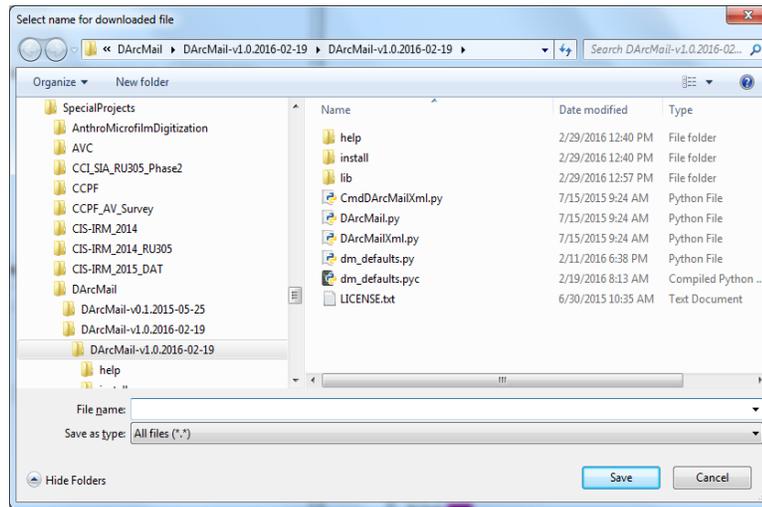
**Original Name.** DArcMail preserves the original names of attachments.

**Part Id.** This is the sequential part number assigned in the original MBOX format input. If there is numbering gap in the Part Id column, then the missing part was a multipart "container", or it was a part with zero length.

When a message part is viewed, it will appear in another numbered tab and appear as indicated below.



If the action specified in the message content list is “download,” you will be prompted to select a location.



## Searching for Addresses or Names

An email address is a string like "john.smith@jamestowncolony.com." An email name is a string like "Smith, John" which might appear paired with an email address in a message header, like this:

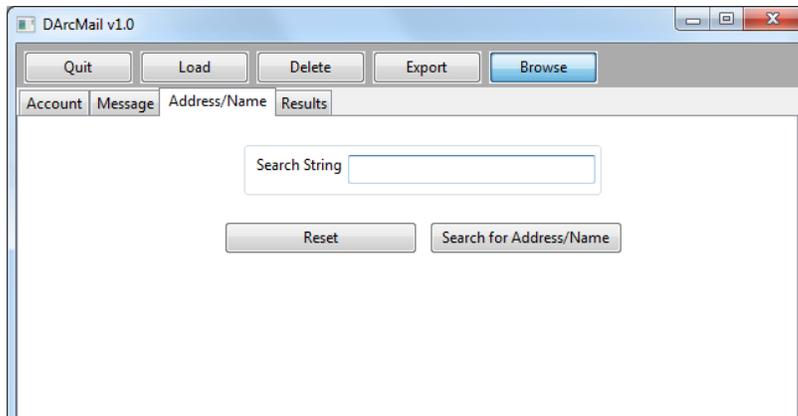
To: Smith, John <john.smith@majestowncolony.com>

Email names are supplied by contact directories, which may differ from user to user, so it is possible for a unique email address to be associated with multiple email names.

To: Cristoforo Colombo <columbusc@expeditions.es>

To: Columbus, Christopher <columbusc@expeditions.es>

Note: It is possible that an MBOX export of a message may have an email name that is not paired with an email address.



### Searching for an email address or email name

- Before searching for an email address or email name, you must set a current account, using the "Set Account" button on an account info page.
- You can search on a partial or complete email address or email name.
- The search string you supply will be automatically wild-carded at the beginning and end.
- Searches are not case sensitive.
- If you do not supply any search string, the search will match all email addresses used in the account.
- If a search string matches an email name but does not match an email address, then the search will return a result only if the email name is paired with an email address in at least one message.

### Address information page

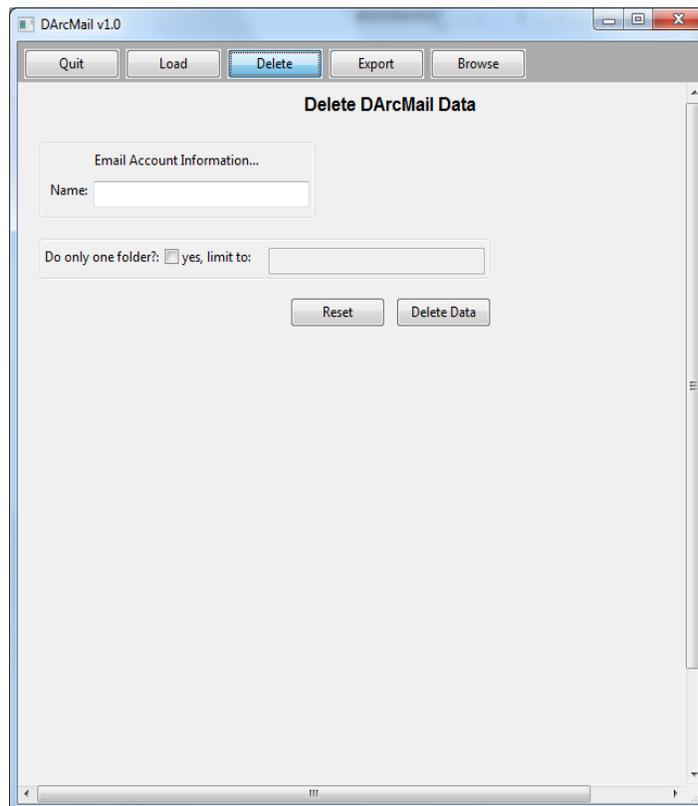
Similar to the message information page, the address information page displays attributes of the email address. Counts of the number of times an email address appears in a From, To, Cc, or Bcc field are linked to corresponding message lists.

## Deleting data from the DArcMail database

You can delete an entire email account from the database, or you can delete a single email folder from an account. In the latter case, this removes that folder's information from the database, as if it had never been loaded. It does not modify the original MBOX file content. If you then export this filtered version of the account from DArcMail, the new MBOX file generated will not include the deleted folder.

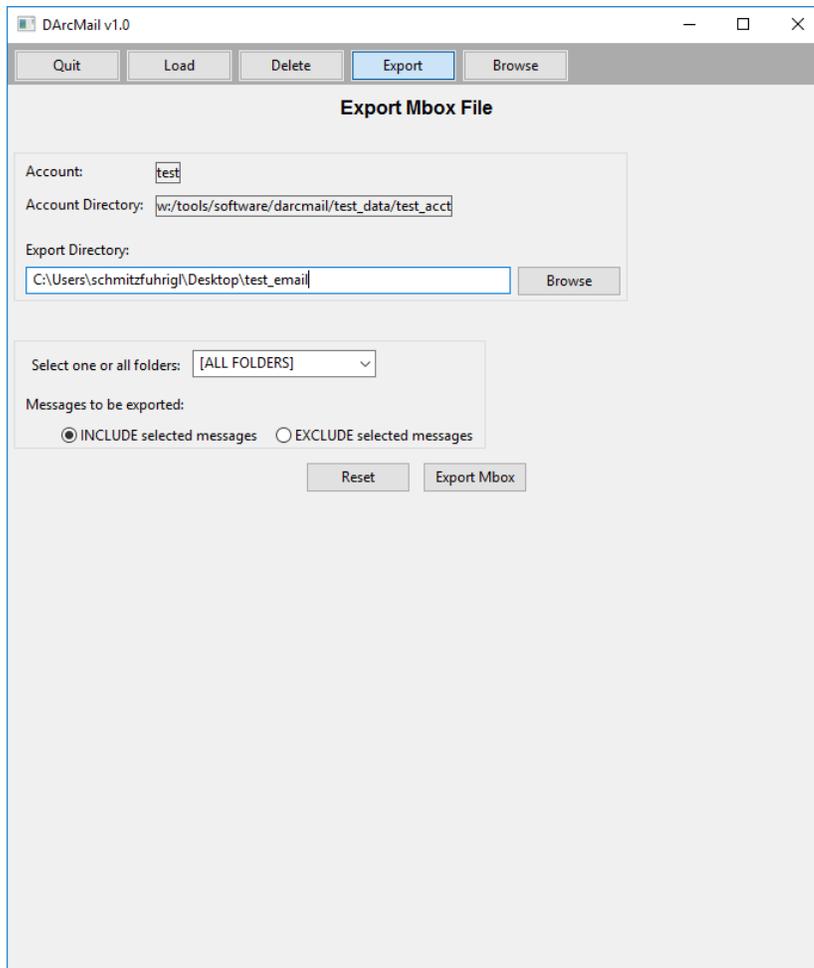
The "Delete DArcMail Data" screen asks for the following information:

- Name. This is the DArcMail name of the email account.
- Folder. If no folder is specified, then the program will delete the email account and all its data from the DArcMail database. If the "yes" box is checked and if a folder "AbcXyz" is specified, then the program will delete only message data that was loaded from the MBOX file named "AbcXyz.mbox."



## Exporting data from the DArcMail database

To export content from DArcMail, choose the Export button.



Specify the location where the exported MBOX file will go under “Export Directory.” Select [ALL FOLDERS] or choose one of the folders in the email account. Indicate if the selected messages are to be included or excluded from the export. Excluding messages may be more useful when the undesired messages are a small number in proportion to the account as a whole.

## Installing DArcMail

### Prerequisites

**DArcMail** runs on Python, using a SQLite3 or a MySQL database.

### Environment

DArcMail can be run from a central or networked environment. In a networked setting, MySQL must be used.

	Server	Clients		
Networked environment	Windows Server 2008 R2, 2012 R2	Windows 7	Mac OSX	Linux
	MySQL Server v 5.5.21 or 5.6.10	Python 2.7.x		
		MySQL/Python Connector v 2.0.4		
		wxPython 3.0.2.x		

Standalone Environment with MySQL	MySQL Server v 5.5.21 or 5.6.10		
	Python 2.7.x		
	MySQL/Python Connector v 2.0.4		
	wxPython 3.0.2.x		

Standalone Environment with SQLite3 (included in Python	Python 2.7.x		
	wxPython 3.0.2.x		

### CmdDArcMailXML

This is the command line version of **DArcMailXML**. It requires Python to be installed.

## DArcMailXML

DArcMailXml allows you to convert one or more .mbox files to an XML file, without using a database. The XML follows the preservation EMA (Email Account) schema, co-developed by North Carolina State Archives and Smithsonian Institution Archives. The schema can be found at [http://www.history.ncdcr.gov/SHRAB/ar/emailpreservation/mail-account/mail-account\\_single.xsd](http://www.history.ncdcr.gov/SHRAB/ar/emailpreservation/mail-account/mail-account_single.xsd).

### Running DArcMailXml

The main DArcMailXml program is the file DArcMailXml.py. If the .py extension has been associated with the Python interpreter (on Windows) or with the Terminal application (on Mac OS X), then you can double-click the file to start the program. Otherwise, you can start the program from a Windows command prompt window or from a Mac OS X Terminal window.

When launched, the "Convert MBOX Files to XML" screen asks for the following information:

**Name.** This is the name of the email account.

**Directory.** This is the account directory. This must already exist and must contain (directly or indirectly) one or more MBOX-format files with the .mbox file extension.

**Folder.** If no folder is specified, then the program will convert all .mbox files in the account directory or in any directory that is a child of the account directory for that account from the DArcMail database. If the "yes" box is checked and if a folder "AbcXyz" is specified, then the program will convert only message data from the MBOX file named "AbcXyz.mbox".

**Max size for internal storage of attachments.** The upper limit, in bytes, of the size of a message attachment if the attachment is to be included in the main XML file. If a message attachment is larger than this upper limit, it will be written to a separate XML file, and a reference to that separate file will be included in the main XML file. If you choose '0' as the max size for internal storage, then ALL message attachments will be stored externally.

**Subdirectories for external storage.** Each email attachment, regardless of size, will be stored in a separate XML file that is referenced in the main XML file. The external XML files for a given folder are stored in the same directory as the .mbox file for the folder. However, if the "Subdirectories for external storage" box is checked, then the attachment

XML files will be distributed among subdirectories of the applicable folder directory. You can use this feature to prevent the creation of directories with thousands of files.

**Split export into chunks.** If "yes" is checked, then the XML will be split into multiple files. The "chunking" is based on message count; you can request that a file contain no more than 1,000 or 5,000 or 10,000 messages.

## XML file names

If you convert an entire email account, then the name of the XML file will be the name of the email account with the ".xml" file extension. If the email account name contains characters that are not permitted in a file name, then each occurrence of an illegal character will be replaced with an underscore. Characters that are replaced with an underscore include any character with a numeric value less than decimal 32 (' ') or greater than decimal 127 ('~'), as well as these characters:

```
/ ? < > \ : * | " ^
```

The XML file for an account will be created in the account directory.

If you export a single folder, then the name of the XML file will be the name of the email folder with the ".xml" file extension. Since an email folder name is the same as the name of an MBOX file (without the ".mbox" extension) all characters in an email folder name are legal in a file name. The XML file for a folder will also be created in the account directory.

If you choose to split an export into chunks, then each chunk will have a "chunk id" between the simple name of the export file and the ".xml" extension, like this:

```
johnsmith.aa.xml
```

```
johnsmith.ab.xml
```

```
...
```

## MessageSummary.csv output

DArcMailXml produces a file named MessageSummary.csv, with one row for each message. The file is in "comma-separated values" so that it can be loaded into a spreadsheet program like Microsoft Excel or OpenOffice Calc. The MessageSummary.csv file has six columns:

```
From
```

```
To
```

Date  
 Subject  
 Message ID  
 SHA1 Hash

If errors are encountered parsing the email messages, the first error message encountered will be noted.

If you choose to "chunk" XML output files, then the MessageSummary.csv will also be "chunked," resulting in file names like Message.aa.csv, Message.ab.csv, etc., where the distinguishing name parts '.aa.', '.ab.', etc. correspond to the chunked XML file names (see above for the naming of XML output files)

Note that the first part of a MessageSummary file is always "MessageSummary," even if it was created for only one folder. If you wish to keep multiple MessageSummary files in the same directory, you must rename them manually.

## CmdDArcMailXml

### Running CmdDArcMailXml

CmdDArcMailXml is a command-line version of DArcMailXml. It has the same options and produces the same outputs as DArcMailXml but does not have a graphical user interface and therefore does not require installation of either MySQL or wxPython.

You can run CmdDArcMailXml from a MacOS X terminal window or from a Windows command prompt window. Command-line options are as follows:

Long Option	Short Option	Kind	Value Type	Meaning	Default
--help	-h			display command options	
--account	-a	Required	string	account name	None
--directory	-d	Required	string	account directory	None

--no_subdirectories	-n	Optional	Boolean flag	do not create subdirectories for externally stored attachments	create subdirectories
--max_internal	-m	Optional	integer	value for maximum internal storage of attachments	store ALL attachments externally
--chunk	-c	Optional	integer	break XML output file into chunks containing N messages in each file	do not chunk output

### CmdDARCMailXML Results

Place in the "Preserved" Directory

XML

MBOX

CSV (subject\_sender\_summary)