ePADD Installation and User Guide

March, 2024

Version 11

ePADD is a software application that supports archival processes around the appraisal, ingest, processing, discovery, and delivery of email archives. It is open source and licensed under an Apache Public License, v2.0.

ePADD development is managed by Stanford University's Department of Special Collections & University Archives, in collaboration with partners at Harvard University and the University of Manchester.

Funding for **ePADD** development is provided through an Andrew W. Mellon Foundation grant. Funding for previous phases of development has been provided by grants from the Institute of Museum & Library Studies (IMLS) National Leadership Grant (NLG) for Libraries and the National Historical Publications and Records Commission (NHPRC).

ePADD

Document Revisions

Date	Version Number	Document Changes
04/21/2014	.01	Initial Draft
04/23/2014	.02	Guide updated to include Appraisal Module
04/29/2014	.03	Guide updated with user comments
09/09/2014	.04	Guide updated with module changes
06/30/2015	1.0	Guide updated for all modules
07/10/2015	1.0.1	Guide updated for all modules reflecting UI updates in 1.0.1 release
08/10/2015	1.1RC3	Guide updated for all modules reflecting new search functionality in 1.1 release
02/01/2016	1.1	Guide updated to reflect IMLS partner feedback
07/15/2016	1.1	Guide updated with announcement of 2.0 beta 1 release
08/19/2016	2.0	Guide updated for all modules reflecting new functionality in 2.0 release
11/21/2016	2.1	Guide updated reflecting new configuration file and other minor changes in 2.1 release
02/27/2017	3.0	Guide updated for all modules reflecting new functionality and UI in 3.0 release
5/22/2017	3.1	Guide updated reflecting minor updates in 3.1 release
7/17/2017	4.0	Guide updated for all modules reflecting new functionality in 4.0 release
8/10/2017	4.0	Guide updates with instructions for standing up Discovery Module on web server
8/15/2017	4.1	Guide updated for all modules reflecting minor fixes in 4.1 release
11/13/2017	4.1	Guide updated to reflect latest OS support
1/22/2018	5.0 beta 1	Guide updated for all modules reflecting new functionality in 5.0b1 release
2/28/2018	5.0	Guide updated for all modules reflecting updates in 5.0 release
5/14/2018	5.1	Guide updated for all modules reflecting small fixes in 5.1 release
6/29/2018	6.0 beta 1	Guide updated for all modules
7/31/2018	6.0	Guide updated for all modules reflecting fixes in 6.0 release
9/10/2018	6.1 beta 1	Guide updated for all modules reflecting small fixes in 6.1 beta release
9/27/2018	6.1	Guide updated for all modules reflecting small fixes in 6.1 release
10/29/2018	7.0 beta 1	Guide updated for all modules reflecting major changes in 7.0 release
12/11/2018	7.0 beta 2	Guide updated for all modules reflecting major changes in 7.0 beta 2 release

1/31/2018	7.0	Guide updated for all modules reflecting major changes in 7.0 final release
7/29/2019	7.1	Guide updated for all modules reflecting changes in 7.1 release
2/20/2020	7.2	Guide updated for all modules reflecting changes in 7.2 release
5/29/2020	7.3 beta	Guide updated to reflect major changes in 7.3 beta release
6/29/2020	7.3	Guide updated for all modules reflecting changes in 7.3 release
1/15/2021	7.3.5 alpha	Guide updated for the Processing Module to reflect changes in the 7.3.5 alpha release
4/8/2021	8	Guide updated for all modules reflecting changes in 8 release
10/5/2022	9	Guide updated for the Appraisal Module and Processing Module to reflect changes in 9 release
1/24/2023	10	Guide updated for the Appraisal and Processing modules to reflect changes in the 10 alpha release
3/1/2024	11	Guide updated for the Appraisal and Processing modules to reflect changes in 11 release

Previous Documentation:

ePADD 10 User Documentation

ePADD 9 User Documentation

ePADD 8 User Documentation

ePADD 6.1 User Documentation

ePADD 5.1 User Documentation

ePADD 4.1 user documentation

ePADD

Table of Contents

1 Overview

2 Getting Started

2.1 System Requirements

2.2 Installing ePADD

2.2.1 Installing ePADD on Windows

2.2.2 Installing ePADD on OSX

2.2.3 Compatibility Note

2.3 Running ePADD

2.4 Selecting a Module

2.5 Navigating ePADD

3 Appraisal Module

3.1 Importing email

3.1.1 Emailchemy Integration

3.1.2 Sidecar Files

3.2 Reviewing Email

3.2.1 Browsing Email Messages

3.2.2 Browsing All Messages

3.2.3 Editing Correspondents

3.2.4 Performing Entity Analysis

3.2.5 Performing Lexicon Analysis

3.2.5.1 Editing a Default Lexicon

3.2.5.2 Reviewing Lexicon Analysis Results

3.2.6 Reviewing Regular Expressions

3.2.7 Reviewing Attachments

3.2.7.1 Reviewing Image Attachments

3.2.7.2 Reviewing All Attachments

3.2.8 Browsing Email Folders

3.3 Appraising Email

3.3.1 Reviewing and Creating Collection Metadata

3.3.2 Searching Email Messages

3.3.2.1 Navigating Search Results

3.3.3 Redacting Email Messages

3.3.4 Annotating Email Messages

3.3.5 Labeling Email Messages

3.3.4.1 Assigning bulk labels

3.3.4.2 Creating a new label

3.3.4.3 Assigning bulk labels by correspondent

3.3.5 Assigning Images to the Collection

3.4 Exporting Email for Preservation

3.5 Exporting Message Attachments

3.6 Exporting Messages to MBOX

3.7 Exporting Email to the Processing Module

4 Processing Module

4.1 Importing an Email Archive from the Appraisal Module into the Processing Module and Assigning Accession Metadata

4.2 Assigning Collection Metadata

4.3 Appraisal Revisited

4.4 Setting Default Labels

4.5 Performing Authority Control

4.5.1 Reconciling Correspondents with Authorized Headings

4.6 Exporting Email for Preservation

4.7 Downloadable .CSV of Search Results

4.8 Exporting Message Headers

4.9 Exporting Message Attachments

4.10 Exporting to MBOX

4.11 Exporting Entities

4.12 Exporting Correspondents

4.13 Exporting Original Text of Non-restricted Messages as Individual Text Files

4.14 Exporting Email to the Delivery and Discovery Modules

5 Discovery Module

5.1 Testing the Discovery Module

5.2 Mounting the Discovery Module under a Web Server

5.3 Searching Entities Across Collections

5.4 Navigating to a Collection

5.5 Browsing Email

5.6 Searching Email

6 Delivery Module

6.1 Importing the Email Archive into the Delivery Module

6.2 Navigating to a Collection

6.3 Browsing the Email Archive

6.4 Searching the Email Archive

6.5 Labeling messages

Appendix

ePADD Data Flows

1 Overview

Welcome to ePADD, the all-in-one email appraisal, processing, discovery, and delivery solution for donors, archival repositories, and researchers. This installation and user guide will assist the new ePADD user in making the most of the program.

ePADD consists of four modules – Appraisal, Processing, Discovery, and Delivery – which are designed to facilitate the process of working with email archives. The application runs in a user-friendly and familiar web browser interface:

Appraisal: Allows creators, dealers, and curators to easily gather and review email archives on a local server (to ensure privacy) prior to transferring those files to an archival repository.

Processing: Provides archivists with the means to arrange and describe email archives on a local server.

Discovery: Provides the tools for repositories to remotely share a redacted view of email archives with users through a public web server discovery environment.

Delivery: Enables archival repositories to provide moderated full-text access to unrestricted email archives within a reading room environment on a local server.

2 Getting Started

This section will get you started working with ePADD from installation through navigating the software platform.

2.1 System Requirements

OS: 64-bit, Windows 7 SP1 / 10, Mac OS X 10.13 / 10.14, Ubuntu 16.04

Memory: > 8 GB RAM

Browser: Chrome 68 or later, Firefox 59 or later.

Java: Java Runtime Environment 64-bit, version 11 or later required for ePADD version 10 alpha.

2.2 Installing ePADD

ePADD has been optimized for Windows 10, Mac OS X 10.14, and Ubuntu 16.04 machines.

Follow the instructions below for your operating system.

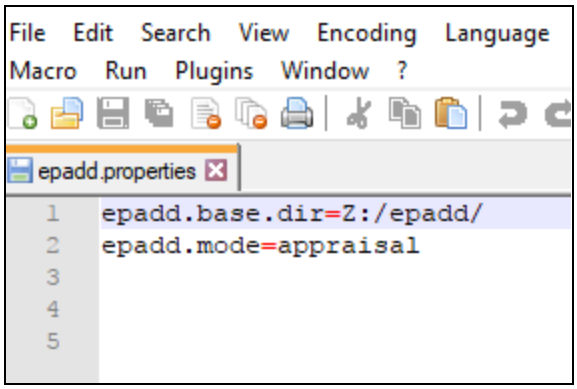
Users are advised to remove previous versions of ePADD before installing a new version. To do so, delete the folder:

On Windows: C:\Users\<username>\epadd-settings
Mac: /Users/<username>/epadd-settings

2.2.1 Installing ePADD on Windows

Please download the latest ePADD distribution files (.exe) from https://github.com/ePADD/epadd/releases/. **You will need to have the Java Runtime Environment version 11 or later installed on your machine for ePADD version 10 alpha to work properly.**

When you run ePADD for the first time, a directory for the Appraisal Module is created to store working files. When ePADD starts up, it checks this directory and relies upon it to resume earlier work. If the software does not locate this directory, ePADD will create it. The ePADD Appraisal Module directory is located at c:\users\<username>\epadd-appraisal.



Example of the custom configuration file

Some settings can be configured using a custom configuration file, *epadd.properties*, which must be in plain text and created in the user's home directory (typically c:\users\<username>\). If the file is not present, ePADD will revert to the existing defaults. Editing this file is only necessary if you wish to change the defaults, for instance, if your default directory does not contain enough space.

1) Using a text editor, edit the existing epadd.properties file.

2) Add to it from the following list of properties.

epadd.settings.dir= The location of the epadd-settings folder. The default is c://users/<username>/epadd-settings. Note the use of the forward slash in the configuration file.

epadd.base.dir= The location of the ePADD module folders (epadd-appraisal, epadd-processing, epadd-discovery, and epadd-delivery). The default is c://users/<username>/epadd-appraisal, etc. Note the use of the forward slash in the configuration file.

epadd.mode= Defines the default module upon opening ePADD. Value should be one of the exact strings, 'appraisal', 'processing', 'discovery', or 'delivery' (without quotes). Match is case insensitive. The default is appraisal.

admin= Defines the admin name and email address to appear on the debug screen. The default is Peter Chan, pchan3@stanford.edu.

3) Save the file.

Depending on your network permissions, you may be asked to allow ePADD access to your internet connection. ePADD requires an internet connection to download email from an email account using the IMAP protocol.

Upon running ePADD, the application icon will appear in the Windows Taskbar. If you do not see the icon, it may be hidden from view. Consult 'Windows help & learning' - https://support.microsoft.com/en-gb/windows - for more information. Right-click on the ePADD icon at any point to open an ePADD window or to quit ePADD.

Note: The Discovery Module is run through a separate distribution file:

Epadd-discovery-standalone.jar

This is accessible via https://github.com/ePADD/epadd/releases/. Be careful not to use the standard ePADD application in a discovery environment. The epadd discovery application should be used in a public discovery environment because it is more secure than the standard ePADD application. For example, it does not allow any modifications to the archive.

ePADD saves all data in folders in the c:\users\<username>\ directory, unless you have defined a different location using the epadd.properties file. We recommend as best practice

making periodic backups of all ePADD related folders in this directory, which may include: epadd-settings, epadd-appraisal, epadd-processing, epadd-discovery, and epadd-delivery.

Note: Depending on your system RAM, ePADD allocates up to 64GB RAM to the application by default. If you wish to run a non-standard amount of RAM, you may wish to run the Java application directly from the command line (epadd-standalone.jar). From the Command Prompt, you can run the application using this command: java -Xmx#g -jar epadd-standalone.jar, where # identifies the amount of RAM (in GB) you wish to allocate.

Note: The user can optionally create and modify a named entity kill list to suppress text strings that ePADD has identified as entities from appearing as such within the software. To do so, create a text file titled 'kill.txt' in the ePADD-settings folder. Each line of the document should contain a single named entity. Save the file and restart ePADD. The entities you listed should be suppressed.

2.2.2 Installing ePADD on OSX

Please download the latest ePADD distribution files (.dmg) from https://github.com/ePADD/epadd/releases/.

When you run ePADD for the first time, a directory for the Appraisal Module is created to store working files. When ePADD starts up, it checks this directory and relies upon it to resume earlier work. If the software does not locate this directory, ePADD will create it. The ePADD Appraisal Module directory is located at Macintosh HD/Users/<username>.

Example of the custom configuration file

Some settings can be configured using a custom configuration file, *epadd.properties*, which must be plain text and created in the user's home directory (typically HD/Users/<username>/). If the file is not present, ePADD will revert to the existing defaults. Editing this file is only necessary if you wish to change the defaults, for instance, if your default directory does not contain enough space.

Follow the following steps to edit the configuration file.

1) Using a text editor, edit the existing epadd.properties file.

2) Add to it from the following list of properties:

epadd.settings.dir= The location of the epadd-settings folder. The default is <HOME>/epadd-settings.

epadd.base.dir= The location of the ePADD module folders (epadd-appraisal, epadd-processing, epadd-discovery, and epadd-delivery). The default is <HOME>/epadd-appraisal, etc.

epadd.mode= Defines the default module upon opening ePADD. Value should be one of the exact strings, 'appraisal', 'processing', 'discovery', or 'delivery' (without quotes). Match is case insensitive. The default is appraisal.

admin= Defines the admin name and email address to appear on the debug screen. The default is Peter Chan, pchan3@stanford.edu.

3) Save the file.

Depending upon your network permissions, you may be asked to allow ePADD access to your internet connection. ePADD requires an internet connection to download email from an email account using the IMAP protocol.

In Mac OSX, the application icon will appear in the OSX Finder Toolbar. Right-click on this icon at any point to open an ePADD window or to quit ePADD.

ePADD requires an internet connection to download email using the IMAP protocol

Note: The Discovery Module is run through a separate distribution file epadd-discovery-standalone.jar, accessible via https://github.com/ePADD/epadd/releases/. Be careful not to use the standard ePADD application in a discovery environment. The epadd discovery application should be used in a public discovery environment because it is more

secure than the standard ePADD application. For example, it does not allow any modifications to the archive.

ePADD saves all data in folders in the Macintosh HD/Users/<username>/ directory. We recommend as best practice making periodic back-ups of all ePADD related folders in this directory, which may include: epadd-settings, epadd-appraisal, epadd-processing, epadd-discovery, and epadd-delivery.

Note: Depending on your system RAM, ePADD allocates up to 64GB RAM to the application by default. If you wish to run a non-standard amount of RAM, or if your installation is running more slowly than you would expect, you may wish to run the Java application directly from the command line (epadd-standalone.jar). From the Command Prompt, you can run the application using this command: java -Xmx#g -jar epadd-standalone.jar, where # identifies the amount of RAM (in GB) you wish to allocate.

Note: The user can optionally create and modify a named entity kill list to suppress text strings that ePADD has identified as entities from appearing as such within the software. To do so, create a text file titled 'kill.txt' in the ePADD-settings folder. Each line of the document should contain a single named entity. Save the file and restart ePADD. The entities you listed should be suppressed.

2.2.3 Compatibility Note

Important: Releases from 6.0 onward package all email archive files using the BagIt specification, and are not backwards compatible with previous releases earlier to v. 5. Users should migrate v. 5 ePADD folders using the script available via https://github.com/ePADD/epadd/releases.

ePADD 5.0 is not compatible with previous versions. Please contact Peter Chan (pchan3@stanford.edu) for assistance in migrating from v4.1 to v5.0.

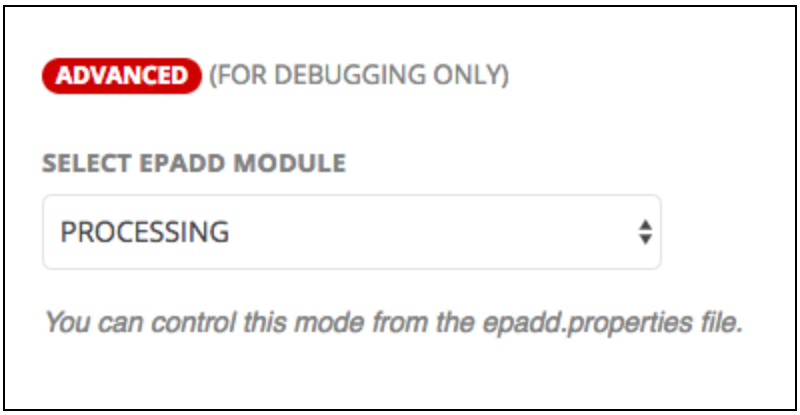
2.3 Running ePADD

To run ePADD, simply double-click the executable file.

2.4 Selecting a Module

The first time you run ePADD, it will default to the Appraisal Module. For testing, this may be changed by manually pointing the browser to localhost:9099/epadd/switch-module. However, for production, it is best practice to first close ePADD, and change the default module using the epadd.properties file (see section 2.2: Installing ePADD in this guide for more details on configuring ePADD).

Note that an email archive must first be ingested into ePADD and exported through the Appraisal Module before it can be accessed in the Processing Module, and exported once again to the Discovery and Delivery Modules before it can be accessed in those modules.



Selecting a module during testing by navigating to localhost:9099/epadd/switch-module

The four modules are:

Appraisal Module – Provides creators, curators, and archivists with a toolset to load, review, and manage an email archive on a local server prior to accessioning that archive to a repository. The module includes various methods to help a user determine the relevance and importance of email messages, to identify and flag sensitive messages, and to impose restrictions on access. Email can be exported from this module to the Processing Module.

Processing Module – Enables an archivist to organize and edit an email archive on a local server following the initial appraisal of materials. Includes the ability to annotate messages, to identify and flag sensitive information, to further refine entity extraction, and to approve/assign authority headings. Email processed through this module can be exported to the Discovery and Delivery Modules. (Note that the Discovery Module is downloaded and run separately; see sections 3.7: Exporting Email to the Processing Module and 5.2: Mounting the Discovery Module Under a Web Server for further information).

Discovery Module - A separate, public web-based discovery environment designed to allow remote users to search email correspondents, headers, and extracted entities, while limiting full-text access to the materials.

Delivery Module – Provides end users with a suite of sophisticated search functionalities to query and review the full content of unrestricted email messages in a collection. This is a client-facing module that is designed to provide mediated access to the processed email archive on a local server running on a managed workstation.

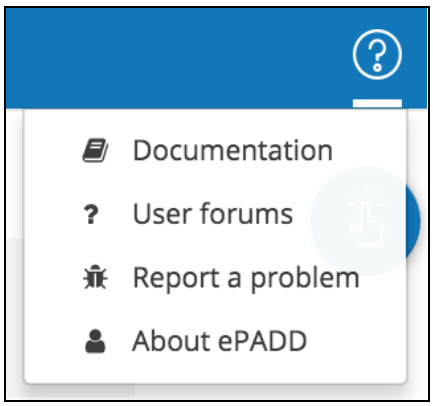
Note: While users can access a Discovery Module from the settings menu, implementing the module from a web server to provide public access requires additional work. Additional information about the Discovery Module can be found in *Section 5*.

Note: We have designed ePADD with the expectation that a given workstation will typically run one specific module at a time, though we recognize this will not always be the case. If you are exploring ePADD for development purposes, and planning to run multiple modules on a single workstation, you may navigate between modules by manually navigating to localhost:9099/epadd/switch-module using the web browser address bar. Please note that it is always best practice to close the archive using the 'Close' header menu option, and to fully quit ePADD, before beginning work in another module.

Please note that it is always best practice to fully quit ePADD and restart ePADD before beginning work in another module. You *must* fully quit and restart ePADD in order to switch to another module from either the Discovery or Delivery module.

2.5 Navigating ePADD

The navigation bar at the top of the screen will present you with primary menu options aligning with the functionality of a particular module. Selecting one of these primary menu options will bring you to a sub-menu offering the potential to carry out related tasks. ePADD is flexible and does not require that the user navigate through a given module in any particular order. Links to documentation, user forums, and other helpful information can be found via the **Help** dropdown located in the far right of the navigation bar.



Documentation and other information can be accessed from the Help dropdown

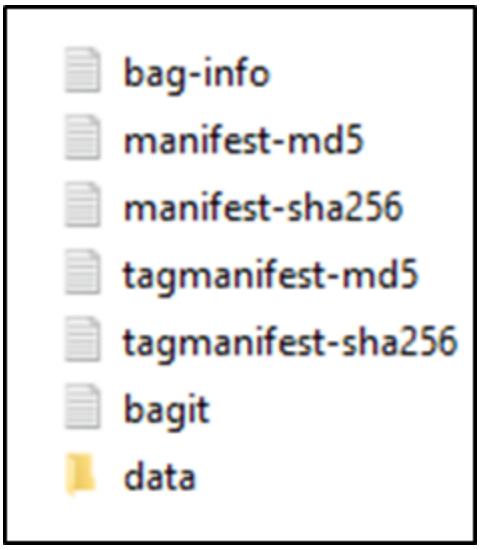
Collections in Processing, Discovery, and Delivery can be closed, and the users return to the collection browse menu, using the 'Close' header menu option, providing additional clarity when working with multiple collections. Note: As always, the user can only work with one collection at a given time.

Additionally, users must manually save their work in Appraisal, Processing, and Delivery, using the 'Save' header menu option.

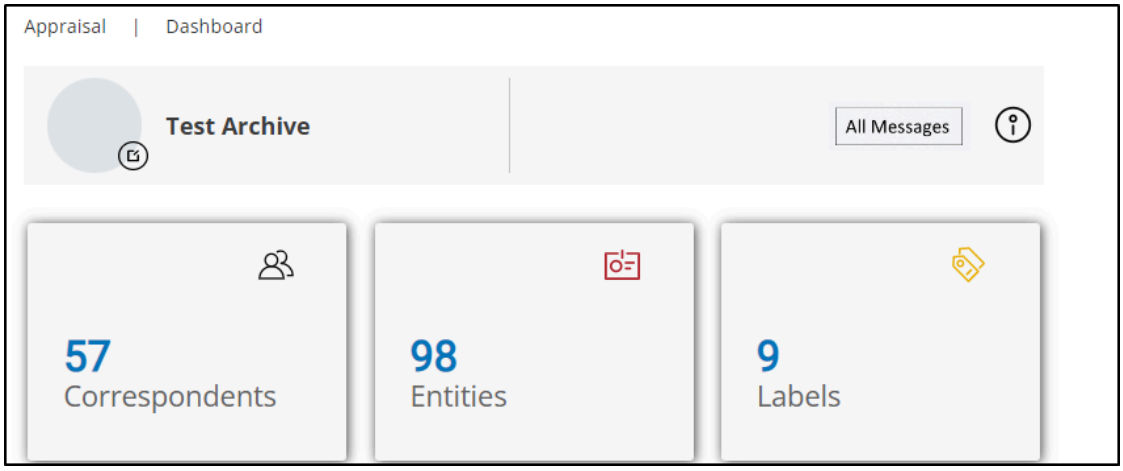
The user can generate a link to a specific message within a given module by selecting the 'Link' option, next to the unique message ID. The generated link appears in the address bar. After loading a collection, pasting the link will take you immediately to that message.

More settings may be available under the **More** option on the Browse screen, including verifying the bag checksum. This function will verify that the checksum of the bag is consistent with ePADD's expectations. This may be especially helpful to ensure that an email archive transferred from a donor has not been corrupted or otherwise tampered with.

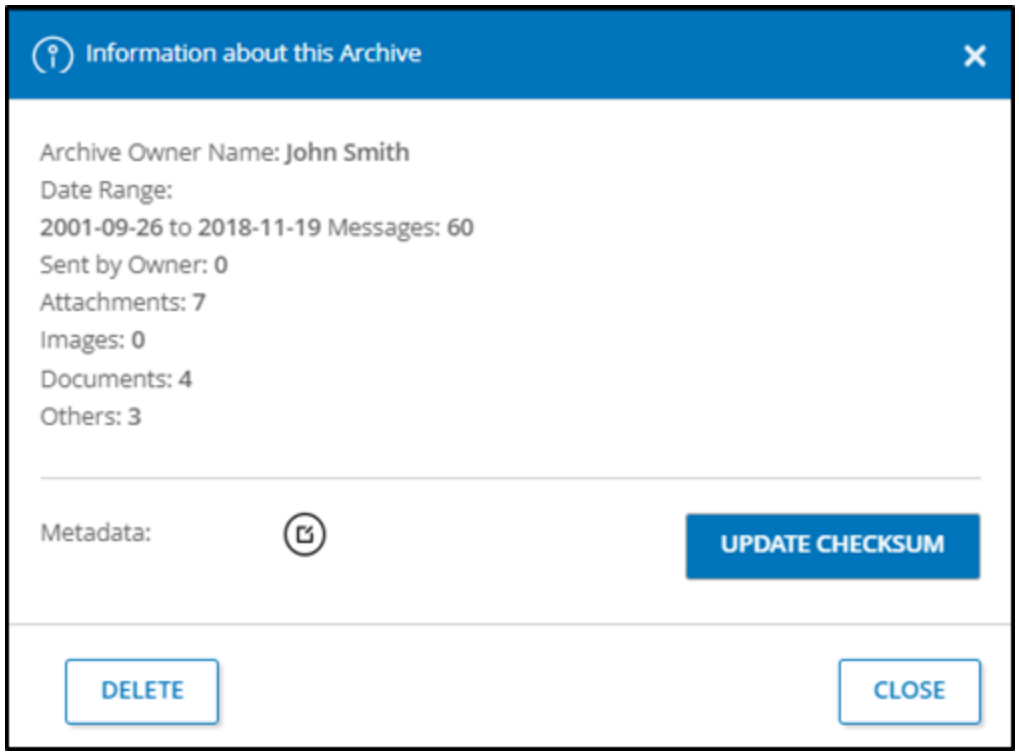
ePADD uses SHA256_fixity checksums and MD5 checksums. Both are present in the ePADD Appraisal directory:



The use of the SHA256_fixity checksums is new to ePADD v.10. If an email archive was imported using an earlier version of ePADD, it will not have been subject to the SHA256_fixity checksum. To apply a SHA256 checksum, click on the Information button next to the email archive title at the head of the ePADD dashboard:

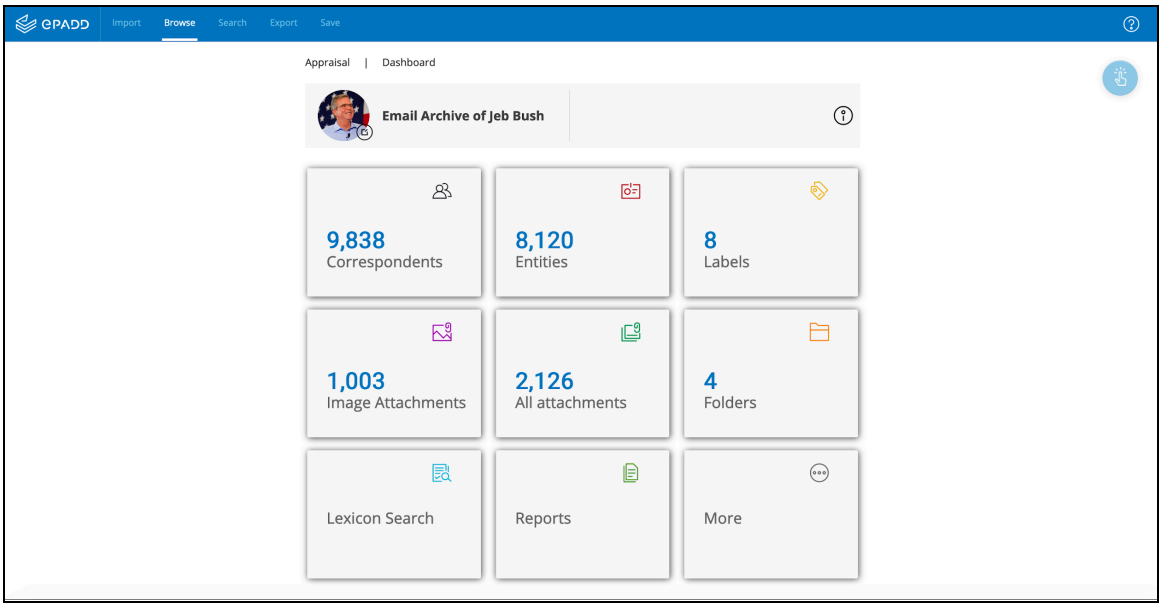


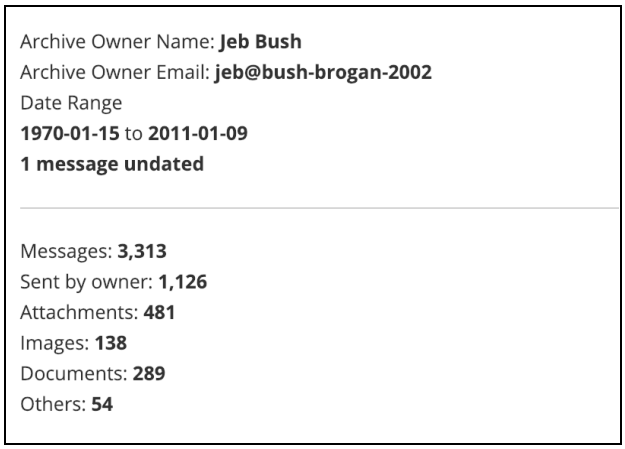
Then, click 'UPDATE CHECKSUM':



We recommend keeping regular backups of your epadd-appraisal, epadd-processing, and/or epadd-delivery folders in case of file corruption.

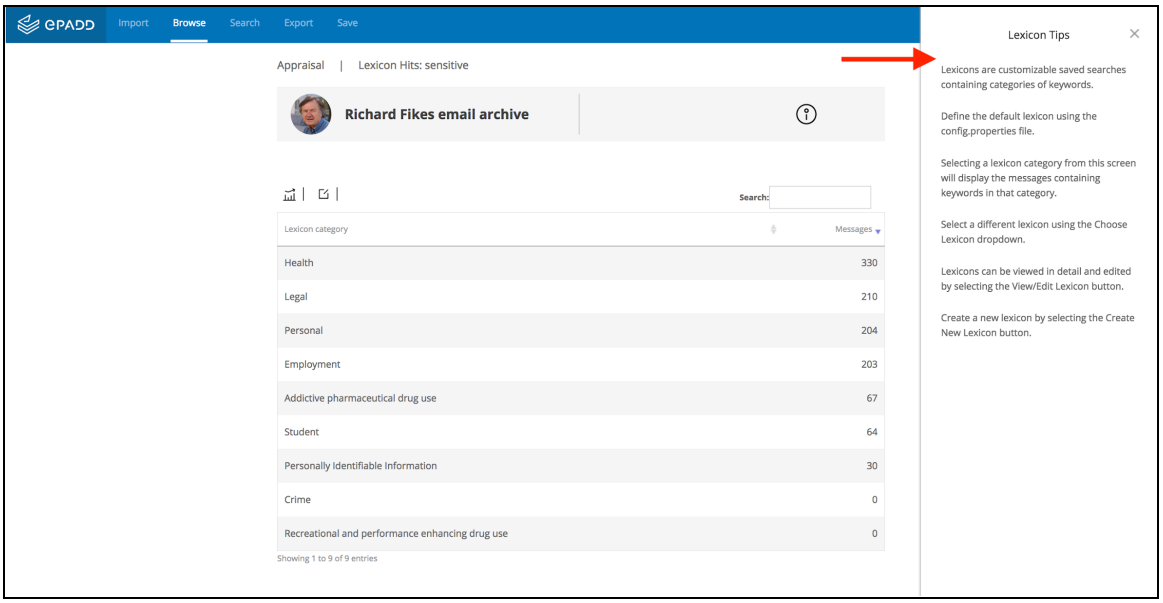
You can view more information about the archive by clicking on the information button adjacent to the email archive title at the top of most pages:





To view additional statistics for the email archive, select the **Reports** option from the Browse menu.

Tips and other helpful information can be accessed on many screens by selecting the help drawer button located in the top right corner of a screen:



An example of a screen which provides helpful information accessed via the help drawer.

3 Appraisal Module

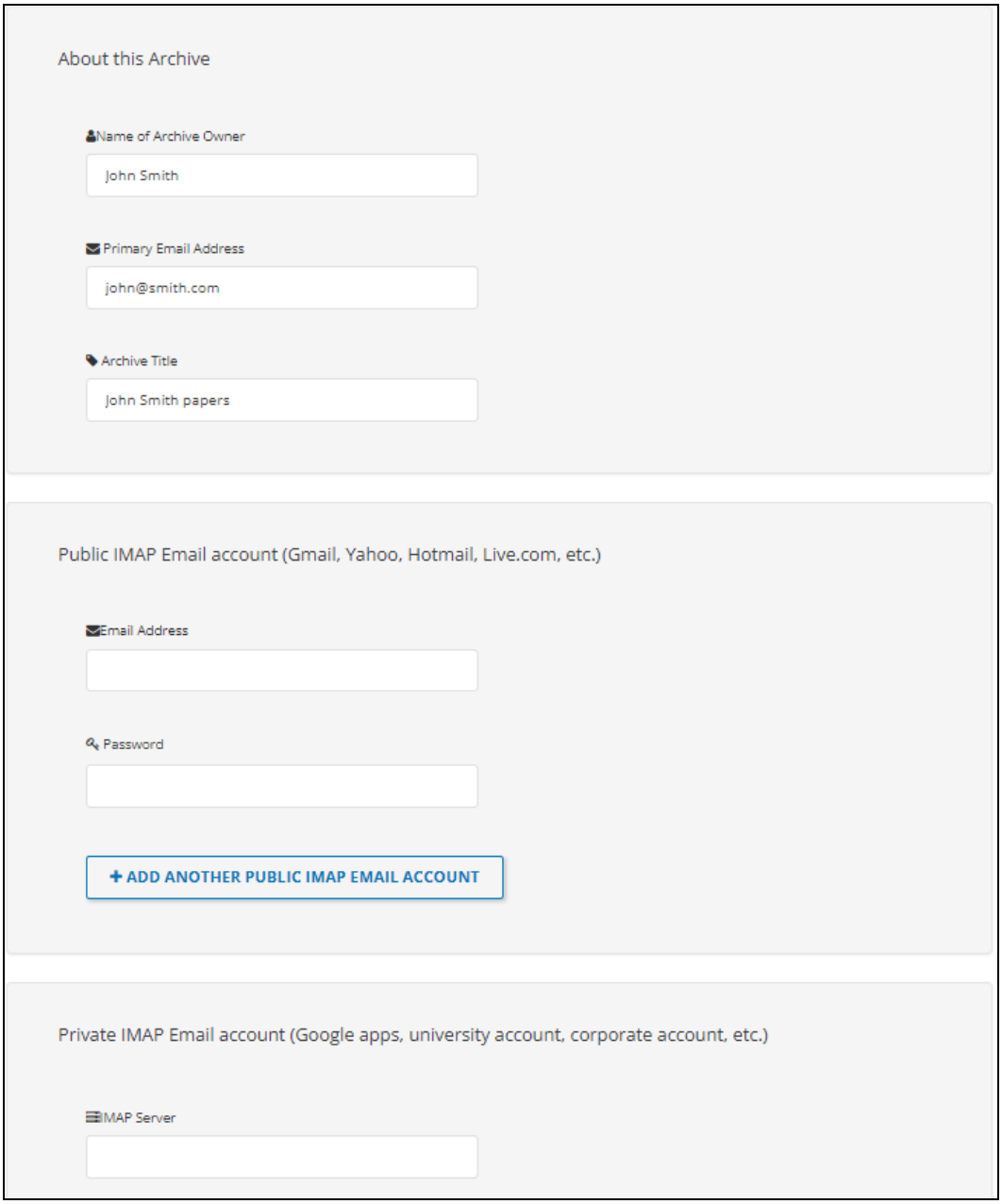
Appraisal Navigation Bar

The **Appraisal Module** provides creators, curators and archivists with a robust toolset to load, review and manage the email of an individual (or optionally, a group of individuals) prior to accessioning that mail to a repository. The module includes various tools to help a user determine the relevance and importance of email messages, to identify and flag sensitive messages, and to impose restrictions on access.

3.1 Importing email

The Appraisal module is the default module (although this can be adjusted using the epadd.properties file). When first launched, you will find yourself on the **Import** screen. This interface is where you enter information (name and an associated email address) about the owner of the email account (optional), as well as specify the location and account info, if applicable, of the MBOX files or IMAP email accounts that ePADD will be ingesting for review and potential transfer to an archival repository. You can select multiple accounts, as well as multiple MBOX files, by selecting the *Add another account* or the *Add* another *folder* buttons.

Note that many email clients do not allow incoming IMAP connections due to their security settings, before attempting an import using IMAP, verify that your account will allow it.

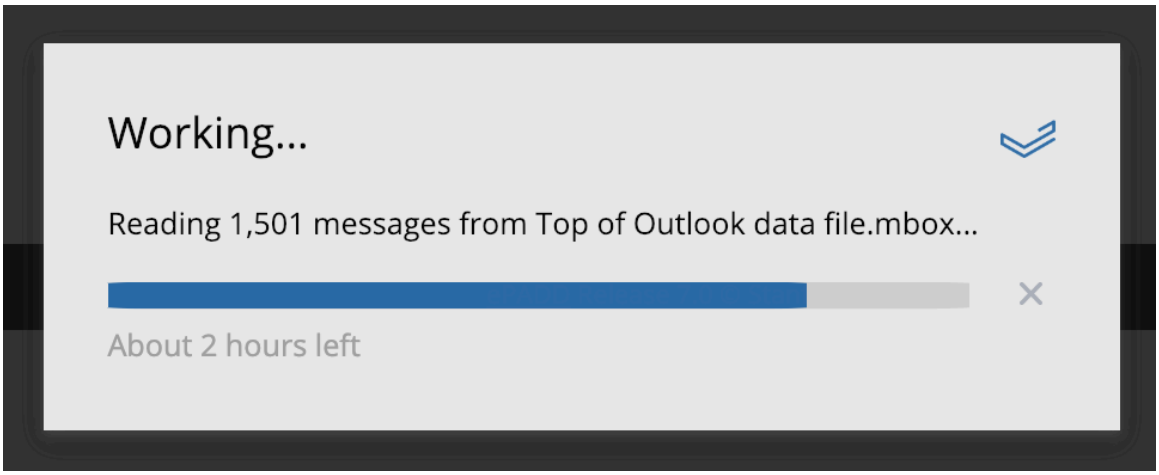


ePADD works with MBOX files as well as IMAP email accounts

Completing the *Name of email source* enables users to explore email by source after import, either by using the new *Folder View*, or by specifying the *Email Source* in Advanced Search. For ingest of email through IMAP, the email account is used as the source name. Please ensure all relevant information has been filled out before pressing the *Continue* button at the bottom of the screen.

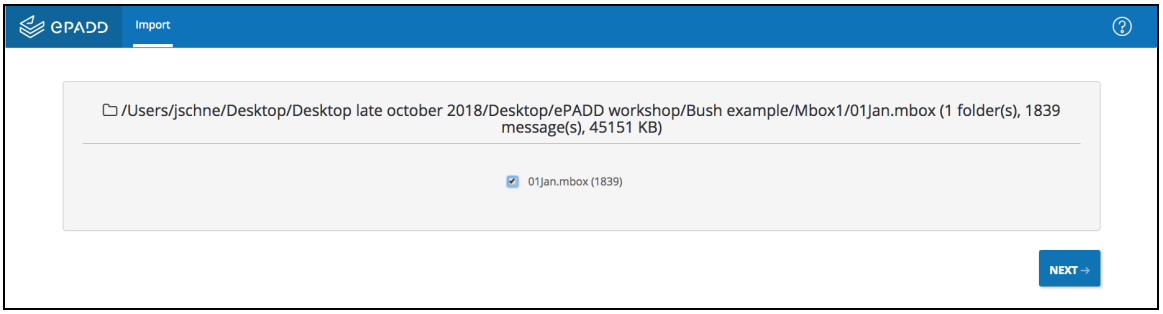
Note: Some IMAP-supported services may require you to confirm through email, text message, or some other means that the login attempt is authentic. Depending on the service, this may require the user to temporarily modify security settings to allow ePADD access.

Additionally, if this is your first time loading these email sources, and especially if you have selected to download mail from an IMAP account, it may take some time for ePADD to perform this initial step.



ePADD will estimate how long it will take to load the selected email files

Once one or more email sources have been loaded, ePADD will give you the opportunity to select or deselect individual email folders associated with accounts, as well as specify a date range to review.



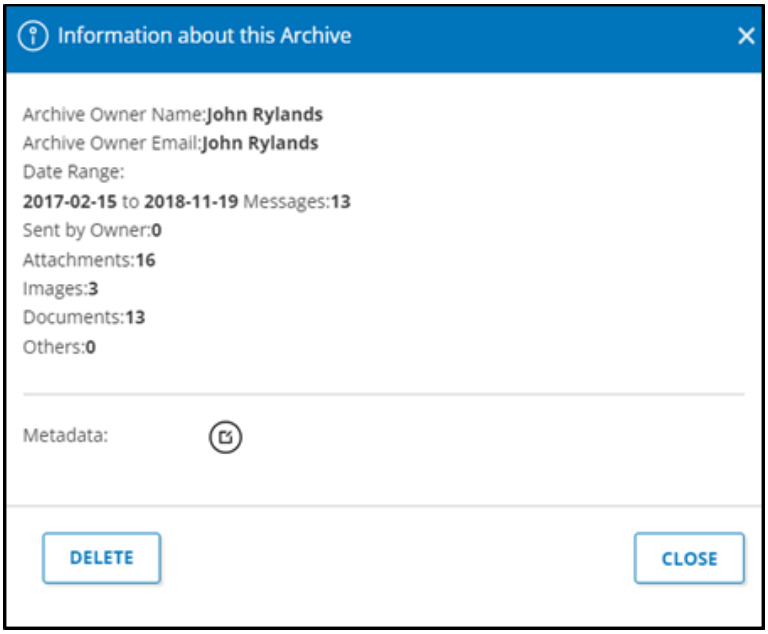
Confirming Selection of Email Sources

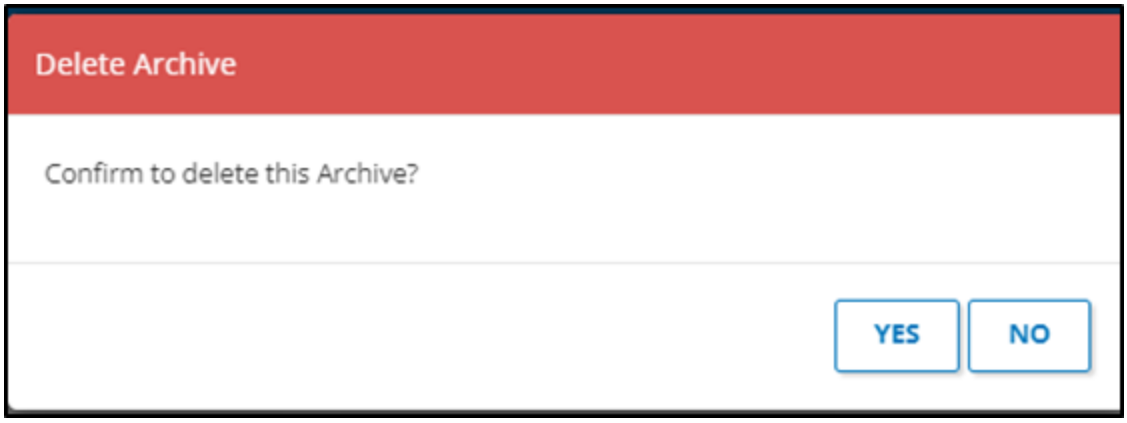
Once you have confirmed the mailboxes you would like to review for transfer, click the *Next* button. Depending on the size of the accounts/folders selected, this next step may take some time as ePADD imports and deduplicates messages and performs a variety of additional steps intended to allow the user rich access to the email archive.

> Note: No further steps are required to export the email archive to an archival repository at this point, if you so choose. To export at this time, proceed to **Section 3.4 Exporting Email**.

If you receive an error during the import process and need to start again, you can delete a collection from within the Appraisal module. Click on the information button - ⓘ - which is located to the right of the email archive's title on the ePADD dashboard:

And then click 'Delete', which is located in the bottom left of the 'Information about this Archive' dialog box:





Alternatively, users can also delete the relevant directory manually by removing the epadd-appraisal folder from the user directory. This folder can be found:

On Windows: C:\Users\<username>\epadd-<module>\
On Mac: /Users/<username>/epadd-<module>\

If you delete the directory manually (not from within ePADD as described below), then make sure ePADD is not running; otherwise, you get a *folder in use* error. You can stop ePADD by going in your browser to localhost:9100.

This method - deleting the directory manually - can be used at any point if you need or want to start again.

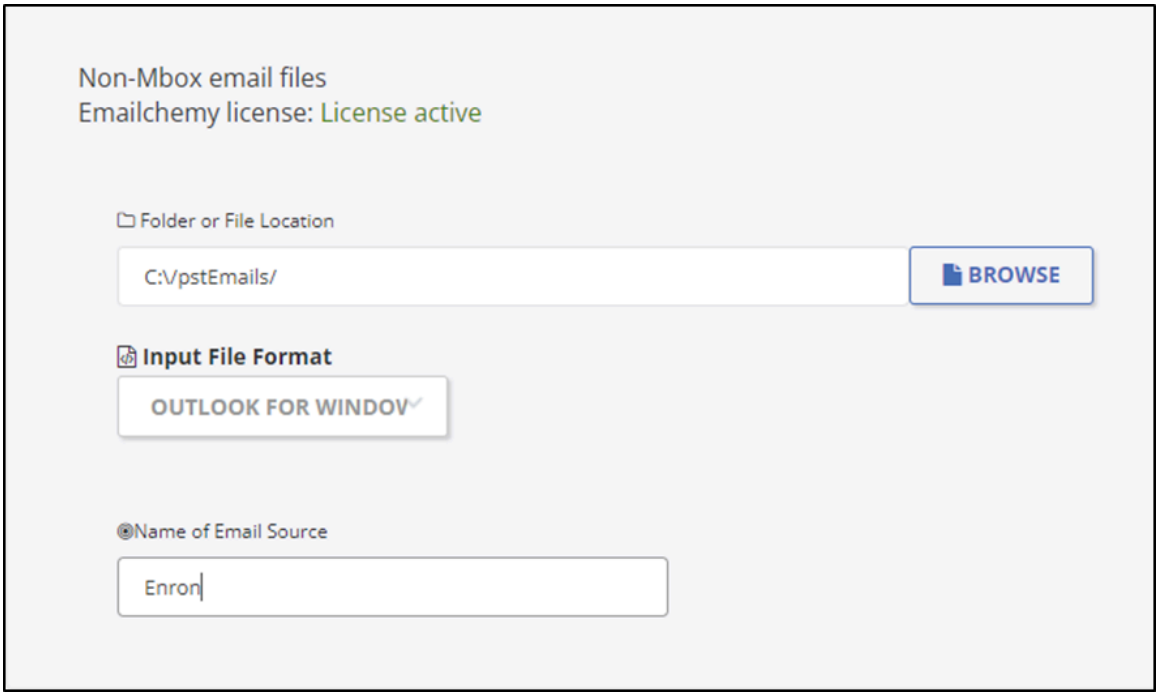
3.1.1 Emailchemy Integration

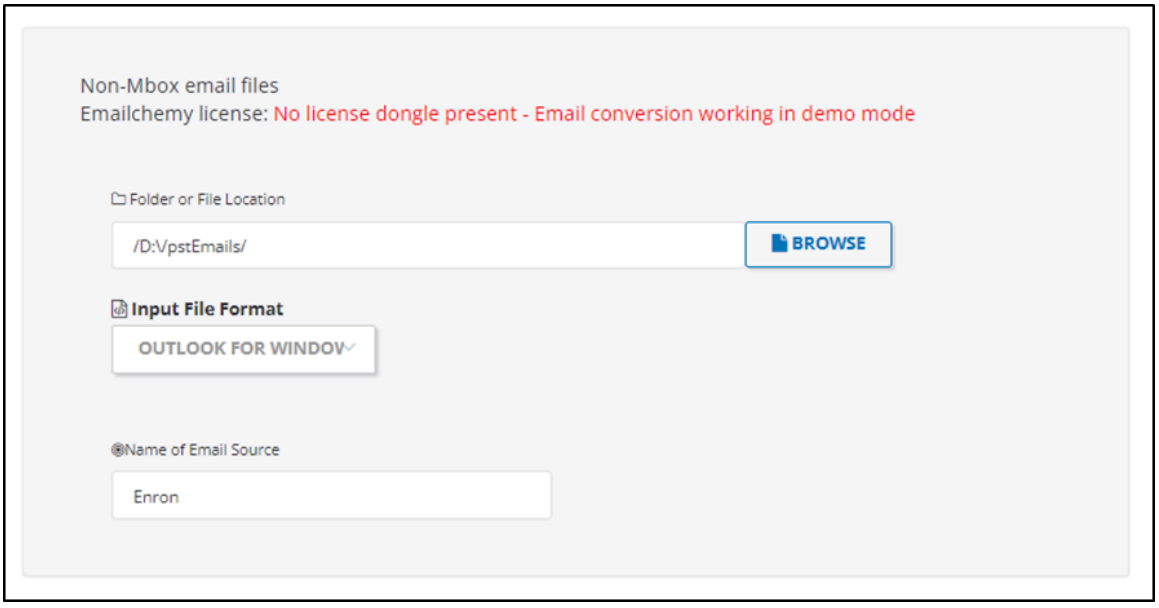
As an optional add-on, ePADD 10.0 offers users the ability to import various email file formats using Emailchemy, an email conversion tool (this add-on is unavailable in earlier versions of ePADD). Emailchemy also allows users to export emails, through ePADD's Appraisal and/or Processing modules, in formats other than MBOX, which is ePADD's standard format. For example, users can export to EML, if that is a preferred preservation format. Click on the following link for a comprehensive list of file formats that Emailchemy reads and writes: https://weirdkid.com/emailchemy/.[1]

To use the integrated version of Emailchemy, users must purchase a license from the developers, Weird Kid Software, and create a license dongle, which will enable those features in ePADD. Further information on how to obtain a license is available via the Weird Kid Software website: https://weirdkid.com/emailchemy-for-epadd/.

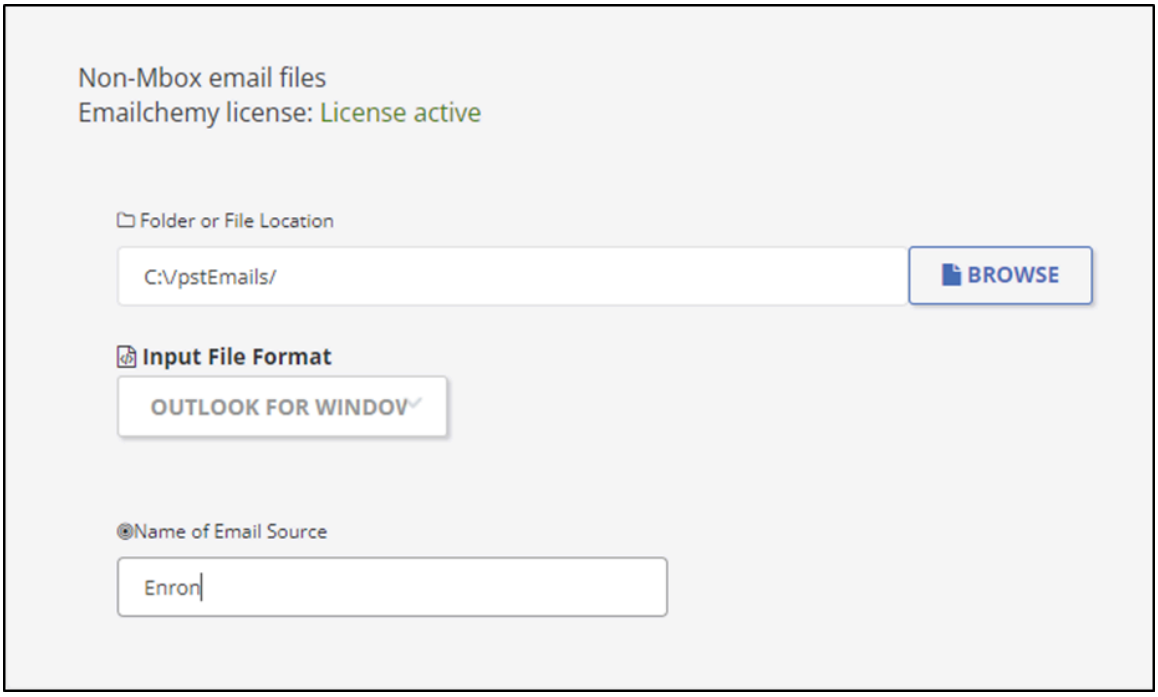
[1] Different versions of the Emailchemy license can write a broader spectrum of formats. This narrower profile of formats was customized for the Emailchemy for ePADD license and selected for their preservation suitability.

The status of the Emailchemy license – whether it has been enabled or not – is noted on the Import screen:

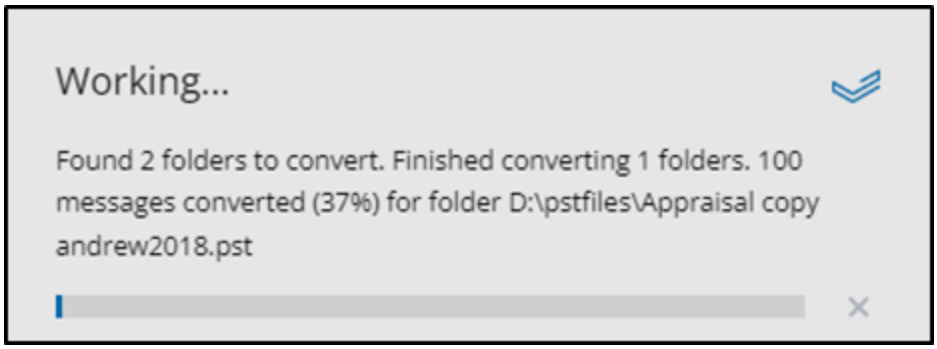




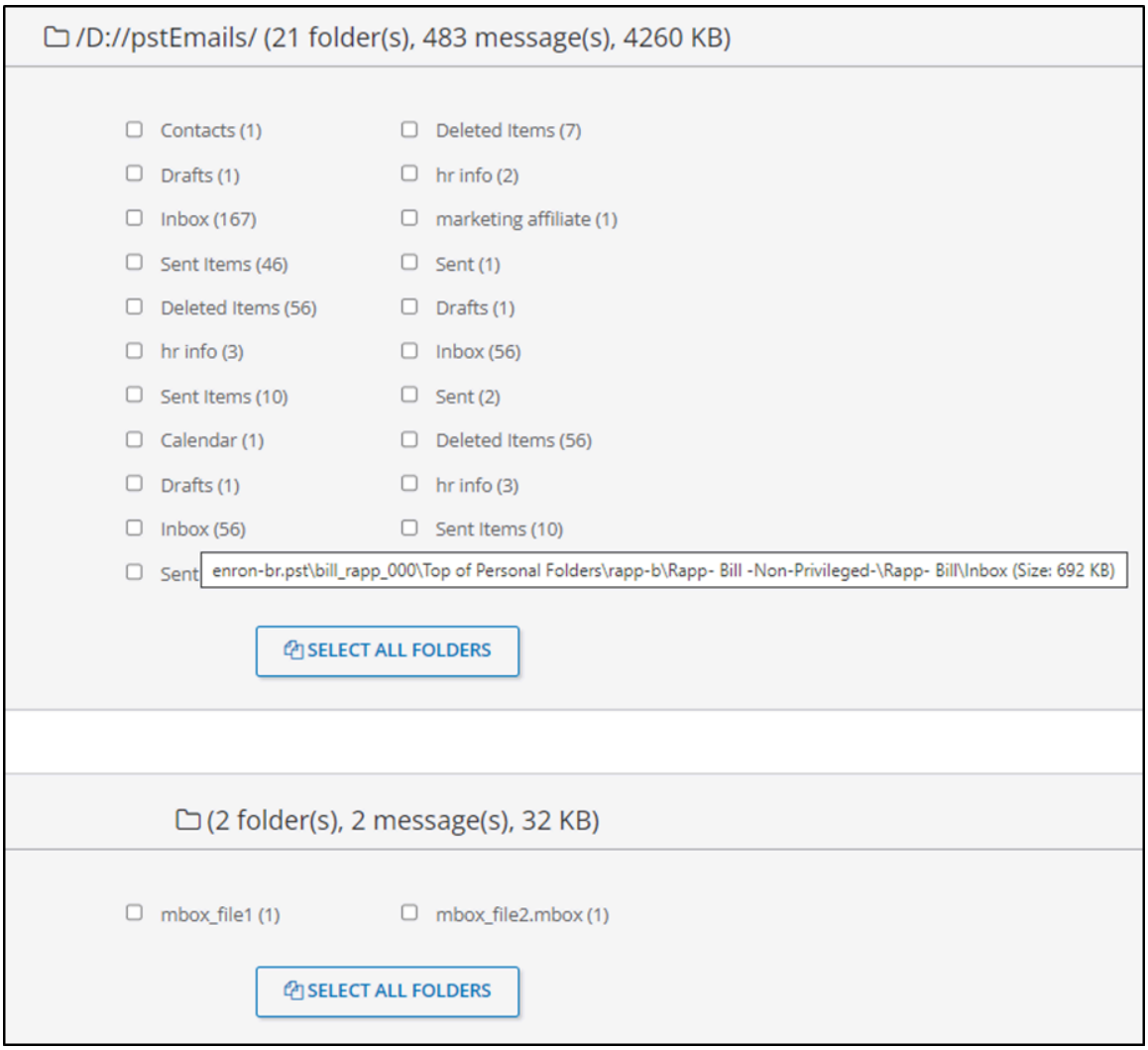
When the Emailchemy license is active, users can point either to a specific non-MBOX file or to a folder containing one or more files:



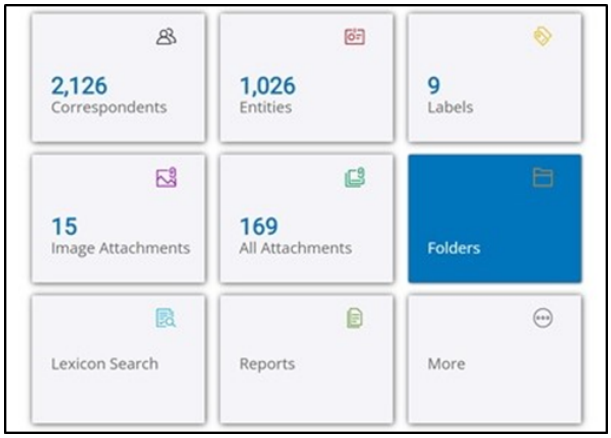
In the example below, two non-MBOX files (PST files) are converted. The term 'folder,' in this context, refers to a non-MBOX file, not the number of folders in that particular file.



ePADD produces a list of folders contained within those non-MBOX files, as in the example below. To see the full file path of each folder, hover the cursor over the folder name. This screen also records the import of any MBOX files.



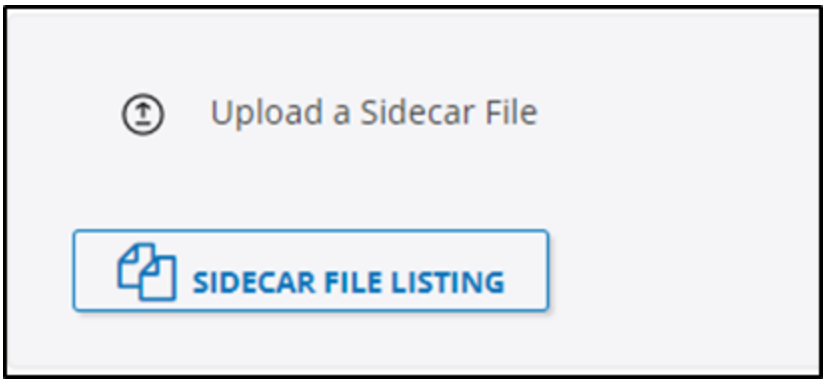
Following the import, select 'Folders' on the Appraisal module dashboard for a list of folders imported from the PST files:



Source	Folder	Messages from Owner	Total
Enron	enron-br.pst\bill_rapp_000\Top of Personal Folders\rapp-b\Rapp- Bill -Non-Privileged-\Rapp- Bill\Inbox	0	56
Enron	enron-br.pst\bill_rapp_000\Top of Personal Folders\rapp-b\Rapp- Bill -Non-Privileged-\Rapp- Bill\Sent	0	2
Archive Kelly	D:\mbox_files\mbox_file1	0	1
Archive Kelly	D:\mbox_files\mbox_file2.mbox	0	1

3.1.2 Sidecar Files

A sidecar file is a file that is not necessarily part of an email archive, but which contains information that is relevant to it, such as a donor agreement, finding aid or metadata file. Sidecar files are optional and can include any type of documentation considered relevant by the user. ePADD allows users to upload, export, list and delete sidecar files. To upload a sidecar file, go to the Import screen and click on the upload icon:



Sidecar files are stored in:

On Windows: C:\Users\<username>\epadd-appraisal\user\data\blobs\Sidecarfiles

On Mac: User/<username>/epadd-appraisal/user/data/ blobs/Sidecarfiles

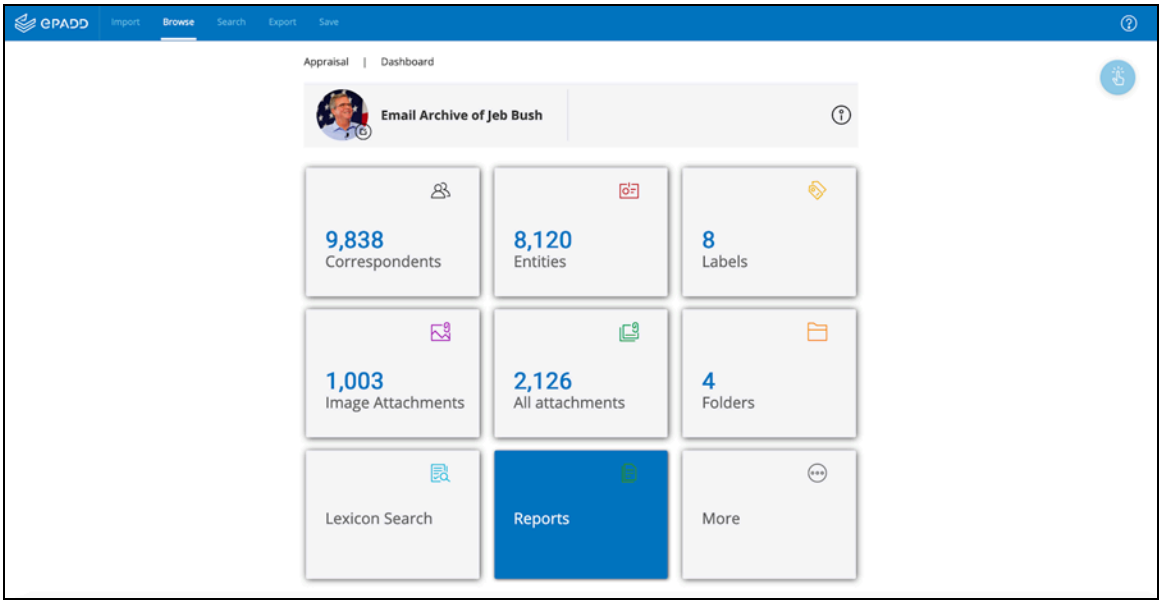
To view a list of uploaded sidecar files and, if necessary, delete any of them, click 'Sidecar File Listing' on the **Import** screen:

3.2 Reviewing Email

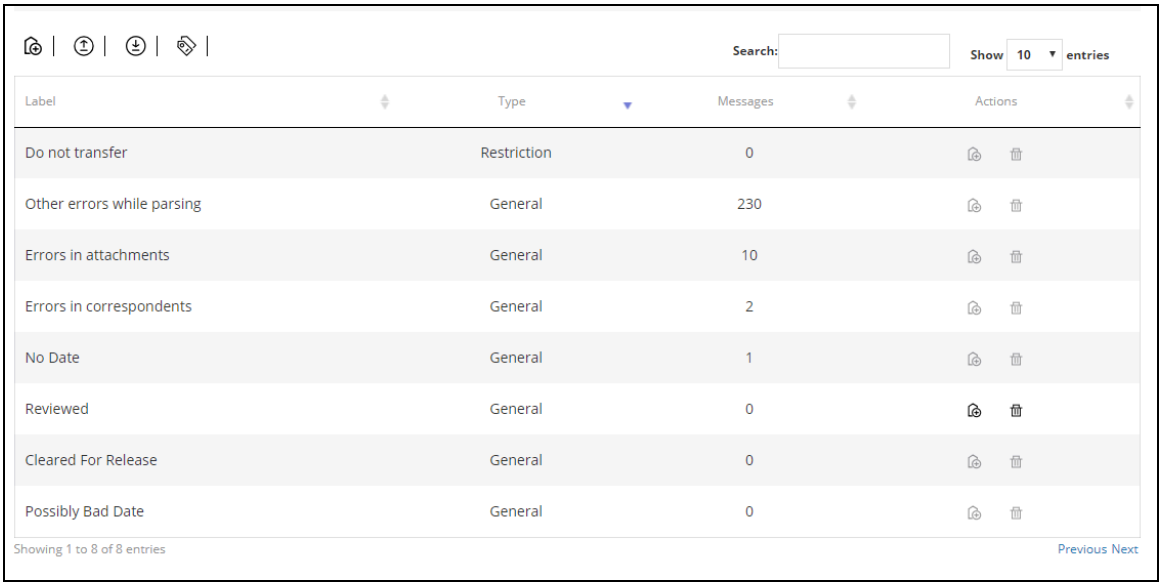
The Appraisal Module provides the user with a diverse set of tools for reviewing and assessing email prior to transfer to an archival repository.

As a first step, the user should review the **Reports** option from the Browse menu.

This screen will present information about the email, including duplicate messages deduplicated by ePADD, and any potential errors identified by ePADD, such as missing headers and dates. Messages containing potential errors are listed on this screen by error type.



The error itself is assigned to the message as a label. To view the set of messages with a particular label, select the Labels option from the Browse screen, then select the label with the associated error type.



Errors may have been generated during creation (in case of older email), transfer, or migration. Ascertaining where the error occurred may require technical expertise or assistance.

With the exception of reviewing the report and potentially investigating possible errors, there is no set order to perform the following steps. In fact, different email archives will benefit from different approaches.

The various tools ePADD provides to assist in email review are described below.

3.2.1 Browsing Email Messages

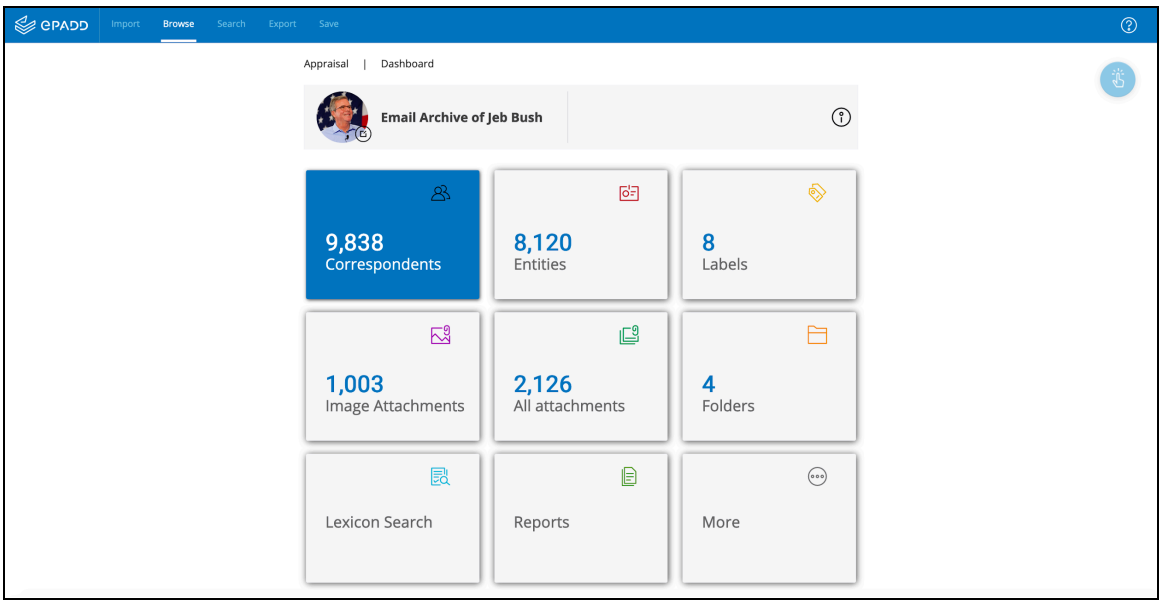
After loading email into ePADD, the program opens the **Browse** screen as the default view. The user can also navigate to the **Browse** screen from the Appraisal Module Menu. From this screen, the user can browse All Messages (which allows users to browse all messages in the collection, one-by-one);by Correspondent; by Person or Other entities; by account folder structure; by Lexicon categories; or by whether messages include Image Attachments, Other Attachments; or contain sensitive information (identified through regular expressions).

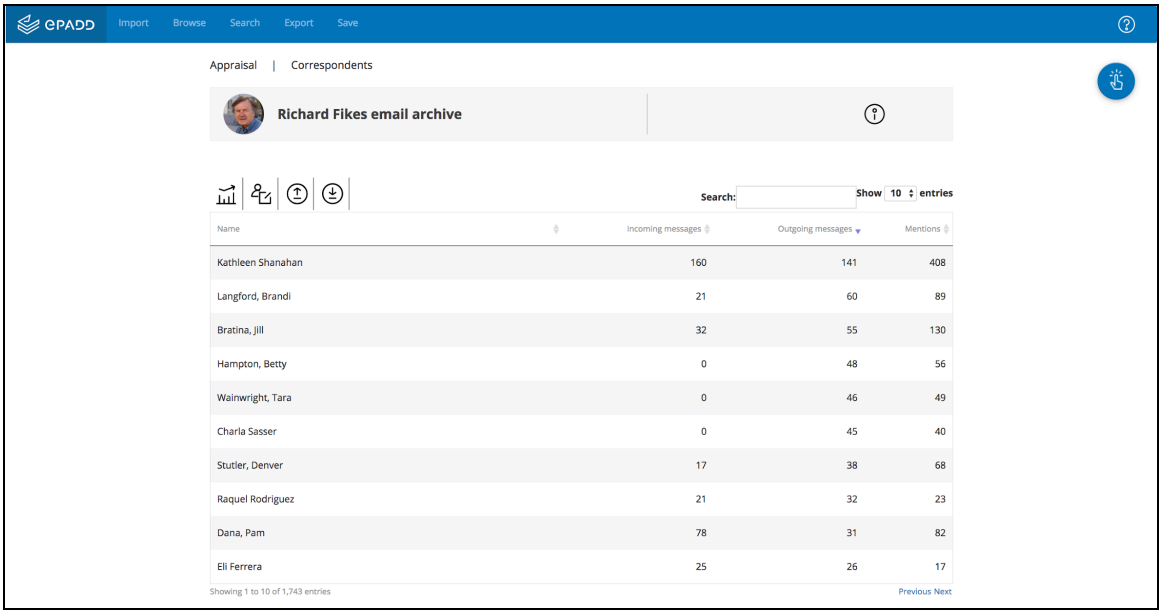
3.2.2 Browsing All Messages

Users can browse all messages in a collection by clicking on 'All Messages', which is located in the banner at the top of the ePADD dashboard:

3.2.3 Editing Correspondents

ePADD automatically merges identities for a single correspondent by intelligently analyzing headers. This step helps clarify relationships between email correspondents to allow for higher-level analysis of the email archives.



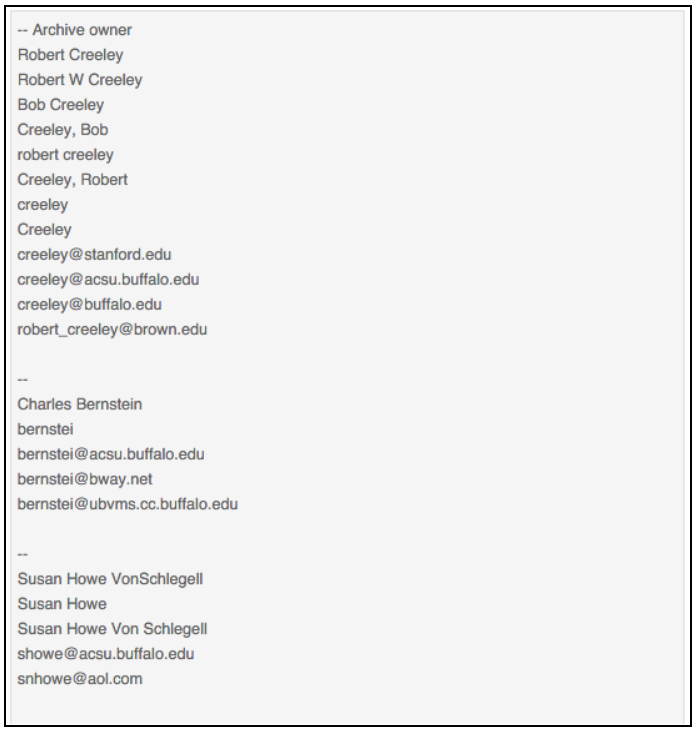


Users can also graph, edit, and export, and re-import the address book using the buttons available from the **Browse Correspondents** screen.

ePADD allows the user to confirm or correct the identities of correspondents that ePADD has resolved through its analysis in order to improve the functioning of the actions that depend upon this behavior.

To do so, navigate to the **Edit Correspondents** screen by selecting the Edit Correspondent button: . This screen allows the user to view and edit the email address(es) and name(s) associated with each correspondent. The default view of this screen lists all correspondents by volume of messages exchanged, but the correspondents can also be listed alphabetically using the dropdown selection at the top of the screen. If there is a single archive owner, they should always be the first individual listed.

Note: As a first step, we recommend downloading the correspondent file before making any edits, as some changes are irreversible.



ePADD allows the user to confirm or correct the identities of correspondents that ePADD has resolved using the Edit Correspondents screen

If names or addresses are incorrectly associated with the wrong individual, you can manually cut and paste them underneath a different entry.

ePADD recognizes header information identifying messages as being sent from a mailing list where that information is available. These messages are differentiated from other correspondents in the Edit Correspondents screen under the heading '-- [ML]' Additional mailing list addresses can be manually identified and added under this heading using the Edit Correspondents interface. Users can also define the default display name for a correspondent when editing correspondents. ePADD interprets the first name listed as the default display name. If you would prefer a different name as the default display, simply add it to the top of the list. **Note: Re-ingesting email will overwrite any user changes. Please be sure to keep track of which addresses have been marked as mailing lists if planning to reingest email.**

When you have completed your work editing correspondents, press *SAVE*.

You can also edit the address book offline, using the export/import functionality available within ePADD.

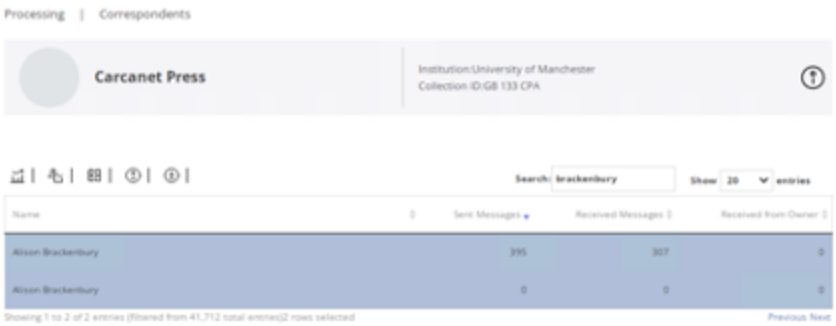
Select the graph button to view a graph of top correspondents. Drag and drop a bounding box over the X axis to zoom in on a time period.

Two or more correspondents can now be selected within the correspondents list in ePADD, and the entries can be combined using a new icon on the correspondents list page in the Processing module.

The entry will now contain all of the email addresses, names and titles contained within the items which were combined.

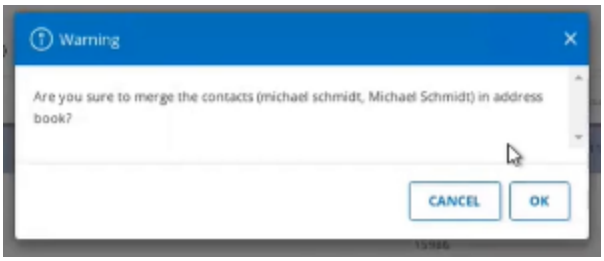
Instructions:

In the correspondents tab, search for the entries you wish to merge and select them by clicking on the entry anywhere other than the name/message figure totals

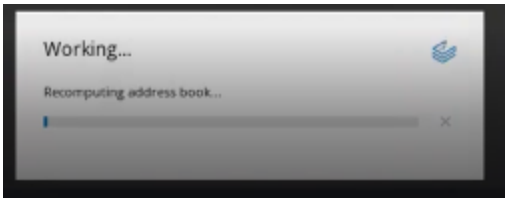


Select the merge correspondents icon:

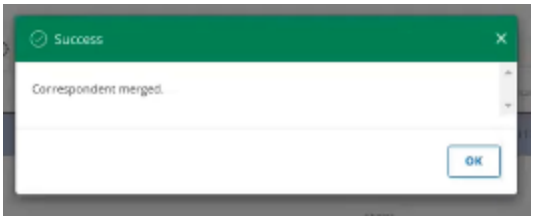
Select 'OK' on the following warning to confirm the merge:



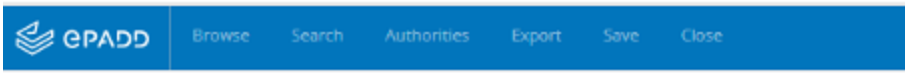
Merging correspondents can sometimes take some time, as it combines multiple entries in the address book but also re-compute contact-related ePADD metadata. A progress bar to provide reassurance that the process is continuing has been added:



Once the merge is complete the following message is displayed:

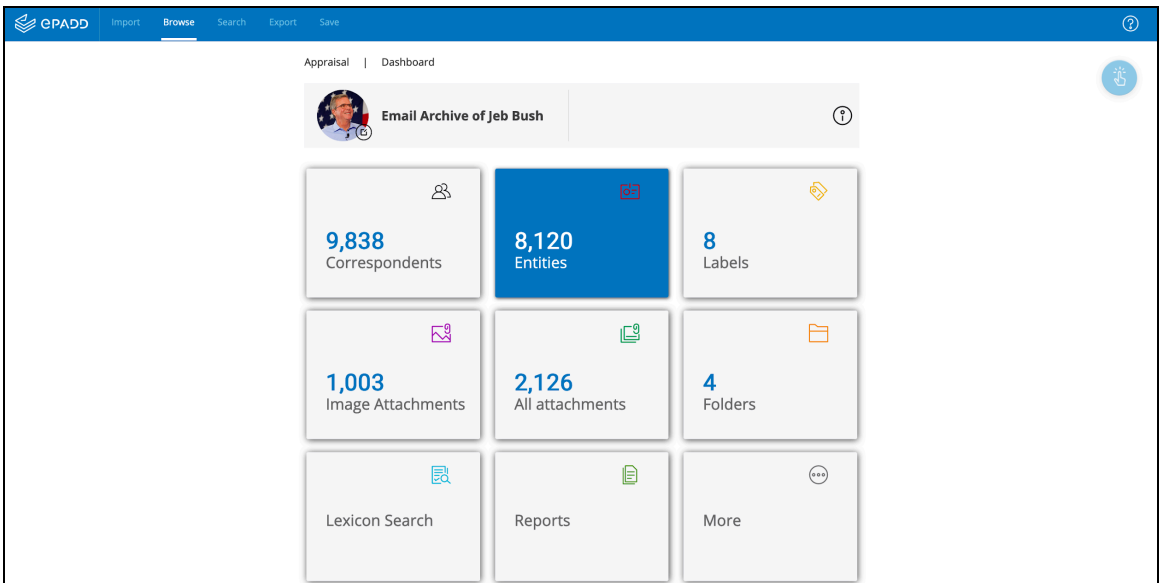


Now select the 'save' option from the top menu, or ePADD will not retain the changes.



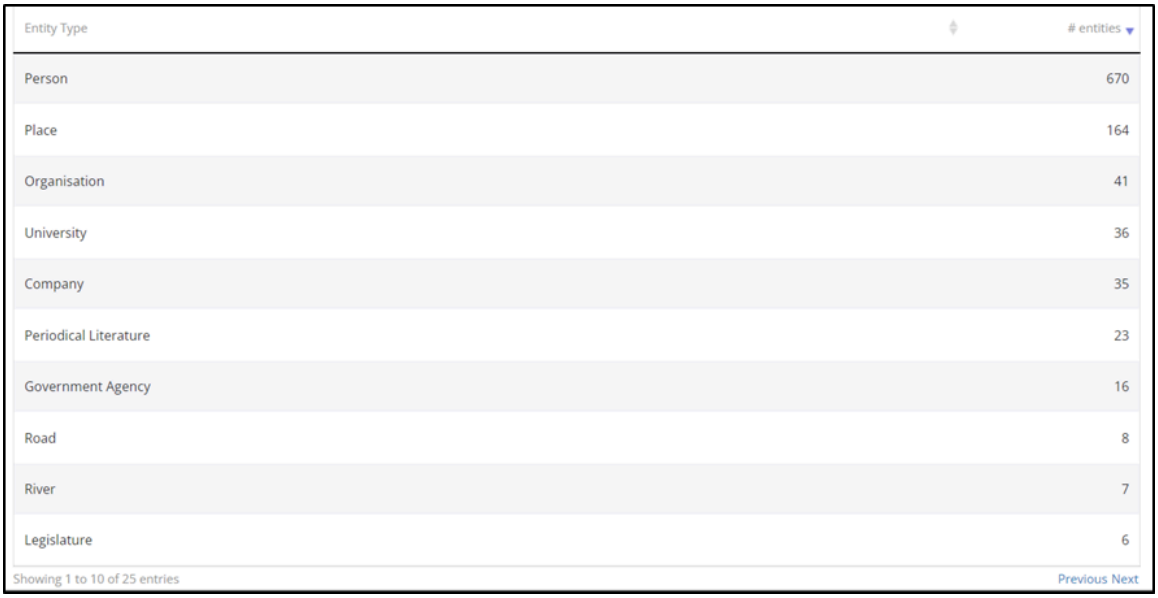
3.2.4 Performing Entity Analysis

ePADD identifies, extracts, and manages entities using a custom natural language processing toolkit. The extracted entities include persons, fine-grained entities such as corporations and government agencies, as well as other entity types such as diseases, awards, and events. Extracted entities can be reviewed by navigating to **Entities**, accessible under the **Browse** menu option from the navigation bar.



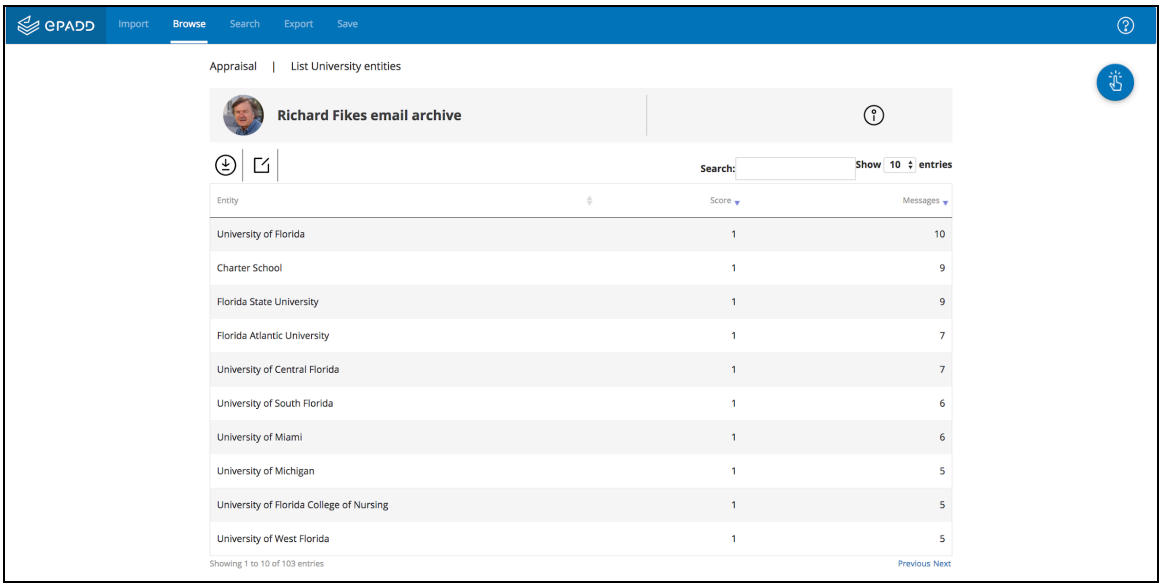
3.2.4.1 Reviewing Fine-Grained Types

To review fine-grained types such as universities and governmental agencies, or other entity types such as diseases and awards, select **Entities** from the Browse menu. You can optionally download a CSV containing all entities and associated metadata (# of messages, score, date range, and entity type) from this screen.



Fine-grained entity types, accessible through the *Entities* option on the browse screen.

Select an entity type to view all of the entities recognized as belonging to that type.

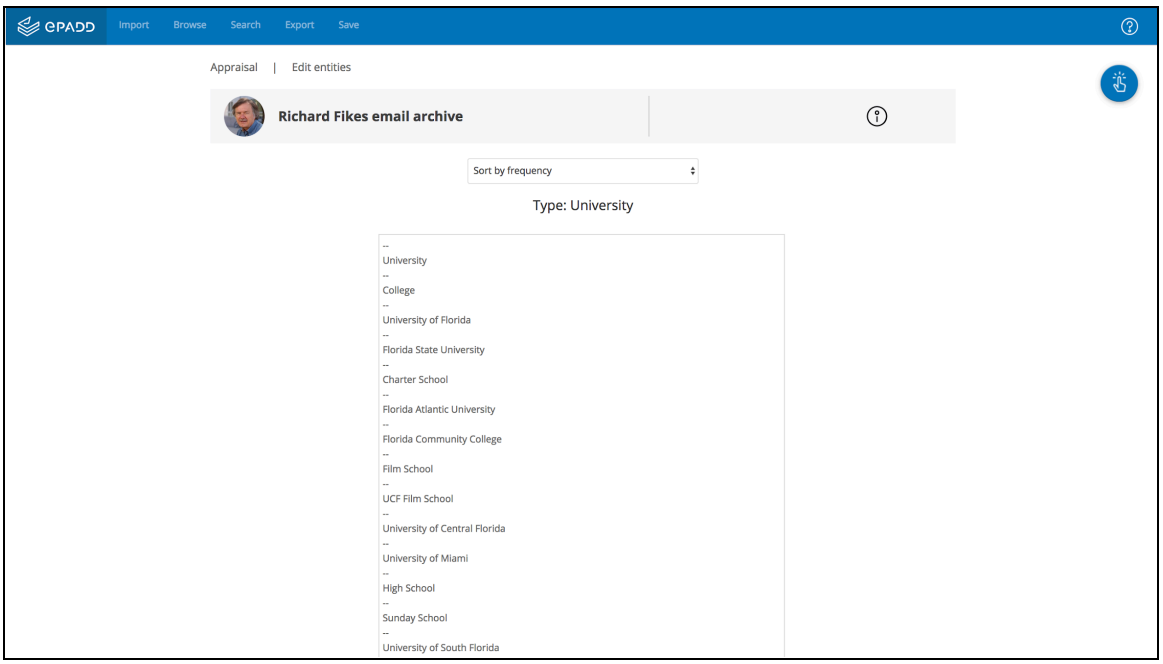


Select an entity to view the set of all messages including that entity. Users can also download all entities of a particular type and associated metadata (# of messages, score, inclusive

message dates) from this screen; or also choose to edit entities associated with a particular entity type.

3.2.4.2 Editing Entities

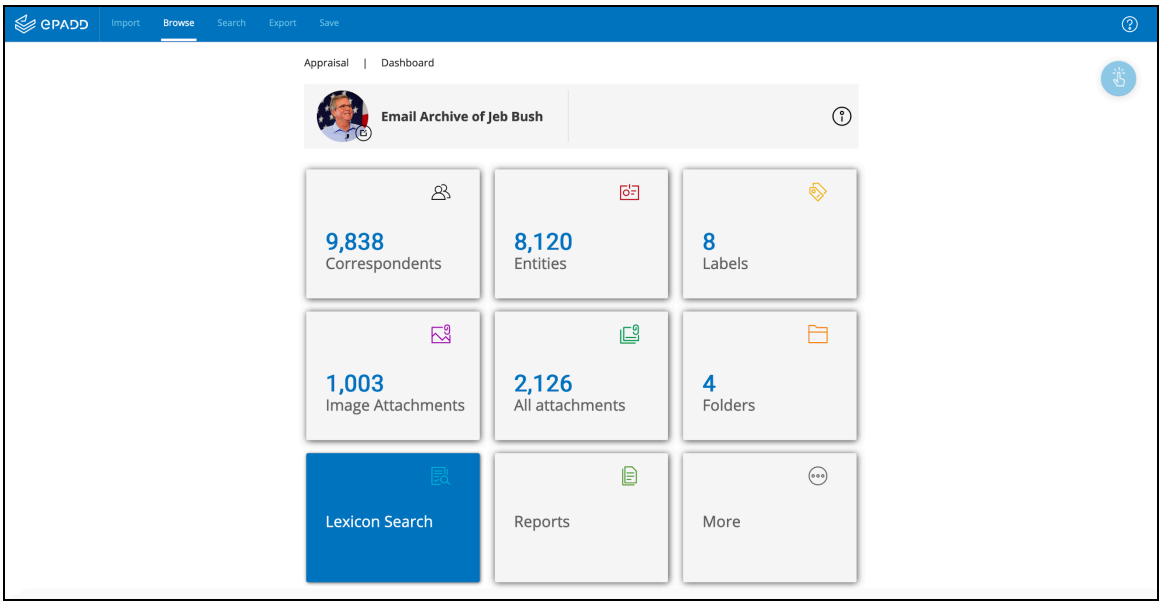
Users can edit or merge entities in a given category from the entity browse screen, by selecting the 'Edit entities' option:



The user can merge entities by grouping them together using find, cut, and paste commands. Unmerge entities by separating them using find, cut, and paste commands. The first entity name listed within each group of entity names is the one ePADD will display in all search and browsing results and visualizations. Manually change this display name by moving a new name to the top of this list. Alternatively, you can supply a new entity display name to the top of the list. This supplied entity name does not need to appear in the email archive. The user can view the individual entities that have been merged with a given entity by hovering over that entity from the Entity Types screen.

The user can optionally create and modify a single entity kill list to suppress text strings that ePADD has identified as entities from appearing as such under any entity category. To do so, create a text file titled 'kill.txt' in the ePADD-settings folder. Each line of the document should contain a single named entity. Save the file and restart ePADD. The entities you listed should be suppressed.

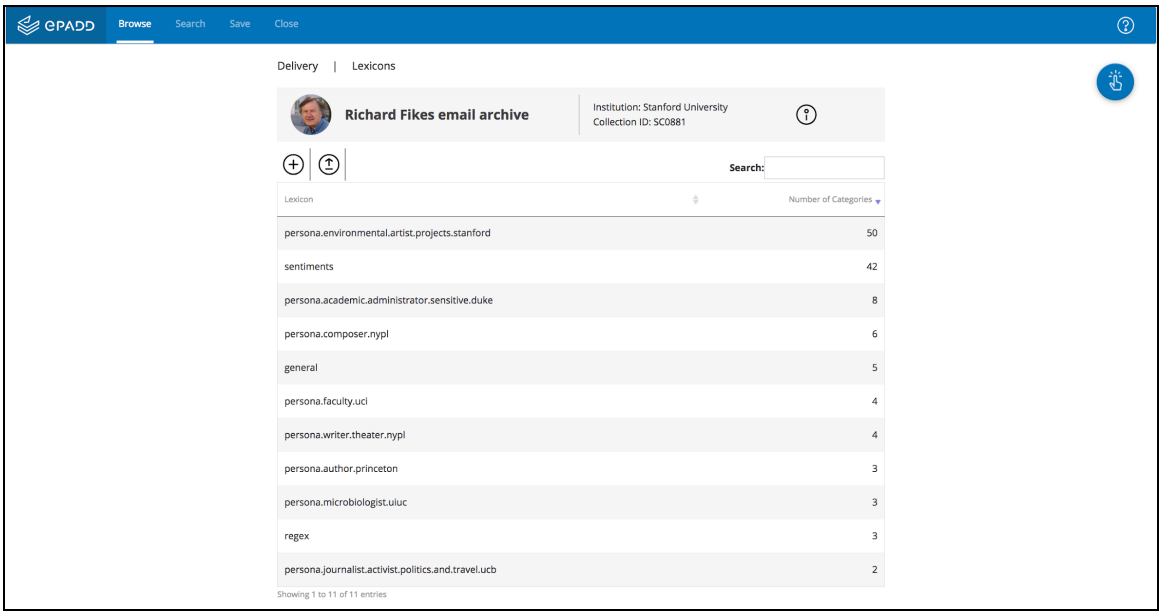
3.2.5 Performing Lexicon Analysis



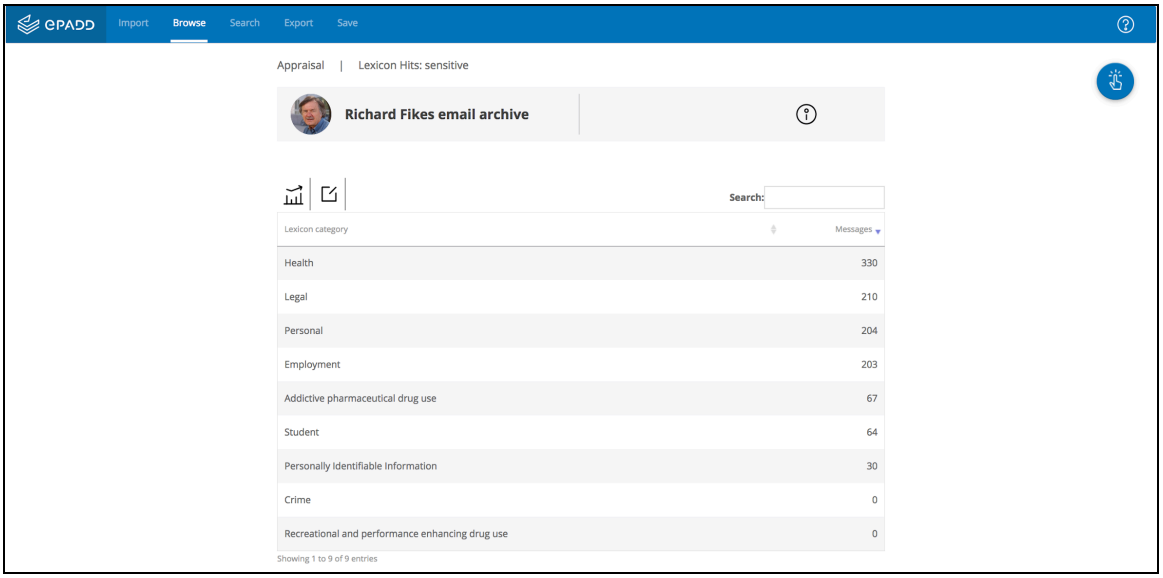
ePADD employs lexicon analysis to search email messages for terms associated with personal or restricted information, which might indicate the need for further review. ePADD ships with several default lexicons. The 'Sensitive' lexicon can be used to assist in the identification of email messages with the potential for confidential content. In release 1.1 ePADD introduced a 'Sentiment' lexicon, to aid the user in performing sentiment analysis on the corpus. Beginning with release 2.0, ePADD includes a variety of additional lexicons developed by a Lexicon Working Group feedback. Additional information about this group, and about the lexicons shared, is accessible here.

The user can make persistent changes to any lexicons, or start afresh by building a new lexicon.

The **Lexicon Browse** screen is accessible under the Browse Menu screen from the Appraisal Module navigation bar.

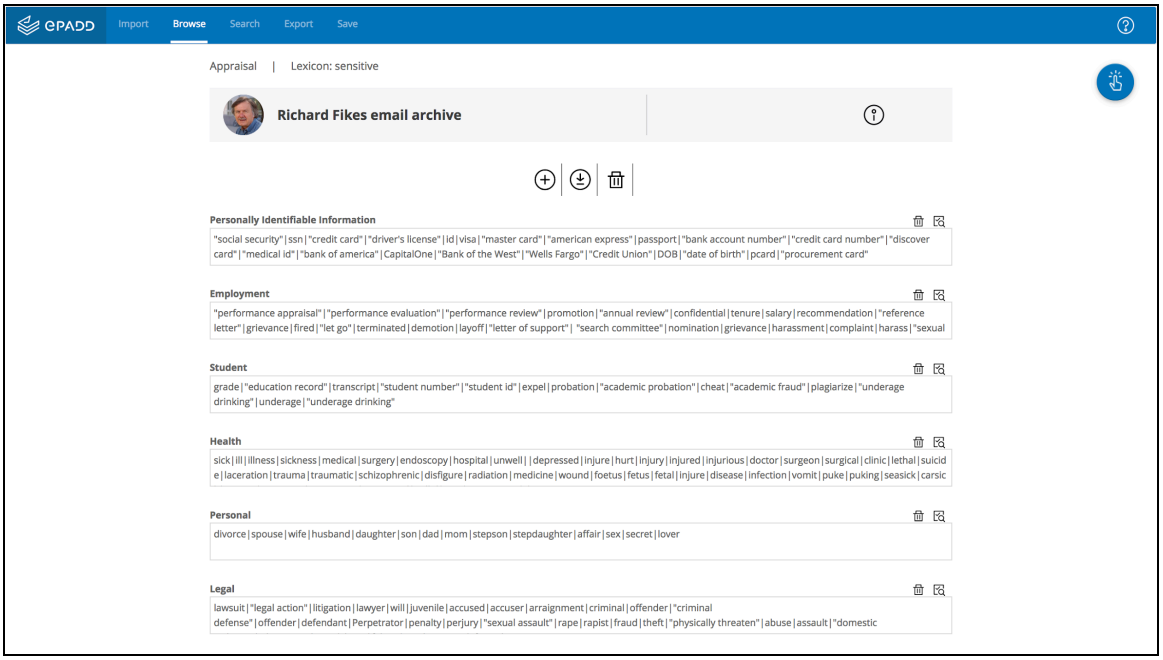


Users can create a new lexicon, import a lexicon from outside ePADD, or view a current lexicon from the Lexicon Browse screen. Click on an existing lexicon to view the included categories of terms and number of hits for each category.



From a particular lexicon page, you can choose to view a graph of the results, or edit the existing lexicon. Or, you can select a category to view all messages that contain a term included in that category.

3.2.5.1 Editing a Default Lexicon



The edit lexicon screen allows users to add categories, delete categories, download the lexicon for offline editing or sharing, or delete the lexicon entirely.

Users can also test categories to view the number of hits for each term in the category. This can be helpful in discovering whether a particular term is too broad or narrow to be helpful.

Please note: New entries in a category should be separated with the | (pipe) character. Multi-word phrases should be enclosed in double-quotes. Note that words and phrases with hyphens, such as father-in-law, should be placed in quotes.

When you have finished making changes, select the *Save* button.

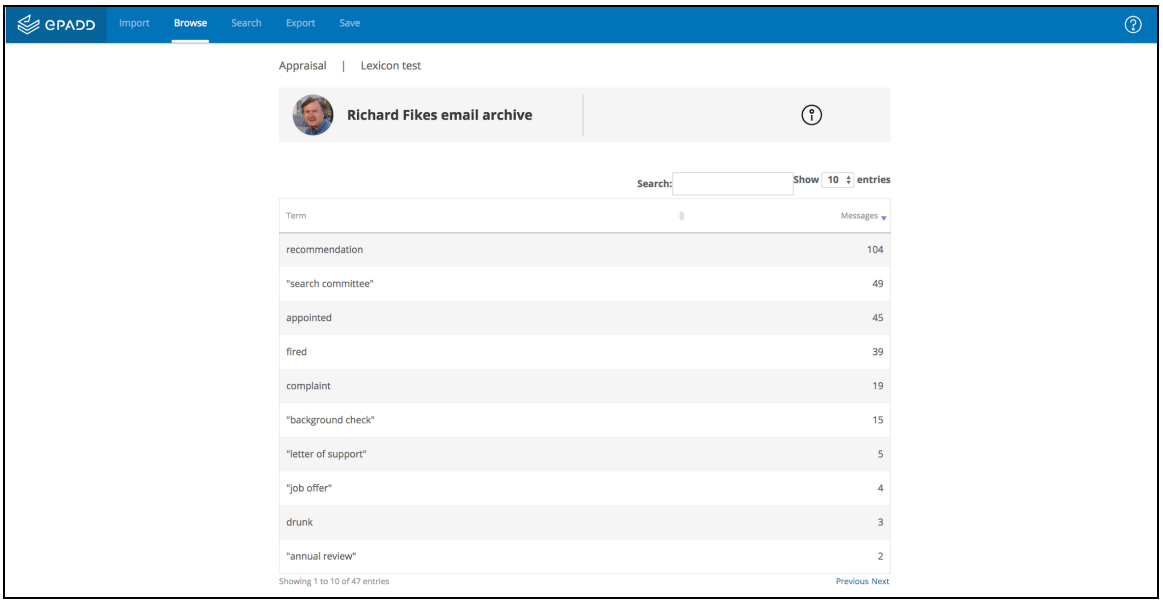
Users can export lexicons using the Export Button, to install them for use with other collections or share them with another institution. In this way, these lexicons can be reused across collections and ePADD instances. Default lexicons may be updated in future releases.

> Expert Tip: The lexicon can also be used to categorize email messages based on topic.

The user **must** use the ePADD interface to add, remove, or update lexicons. This is because of how the bag checksum is calculated.

3.2.5.2 Reviewing Lexicon Analysis Results

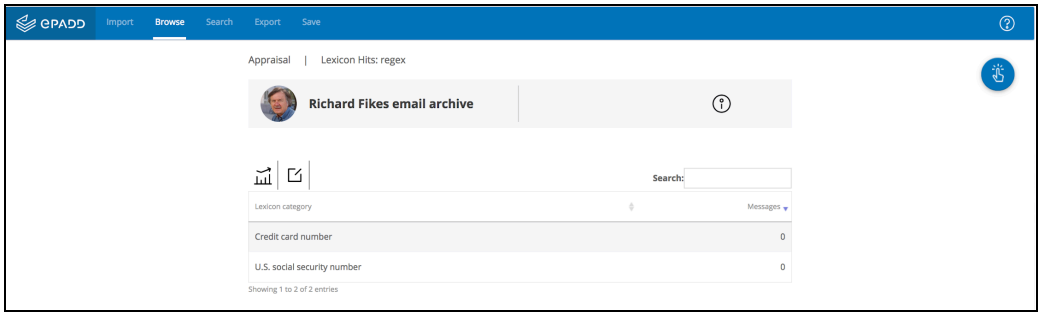
ePADD allows a user to test the results of any category in order to ascertain whether a particular term is too general or too specific to be helpful.



Test terms contained in a lexicon category by selecting the test function from the Edit Lexicon screen

Lexicon analysis results for a whole lexicon may be reviewed via Advanced Search. A user can also view lexicon analysis results directly by clicking category results from the Lexicon Table screen.

3.2.6 Reviewing Regular Expressions



To further aid in identifying potentially confidential materials, ePADD searches the archive for sensitive messages by searching for regular expressions such as social security numbers, and credit card numbers. Regular expressions are a customizable lexicon under the Lexicon Search browse category, selected from the lexicon dropdown menu. Note that since each

lexicon category contains just one regular expression, results for each category display under Lexicon Search results. The user can find the form of many common regular expressions online; https://regexlib.com/ may be a helpful resource in this regard.

3.2.7 Reviewing Attachments

ePADD allows the user to browse and review email attachments. ePADD uses the Apache Tika toolkit to recognize a wide range of document formats. A complete list of formats recognized by apache Tika can be found here.

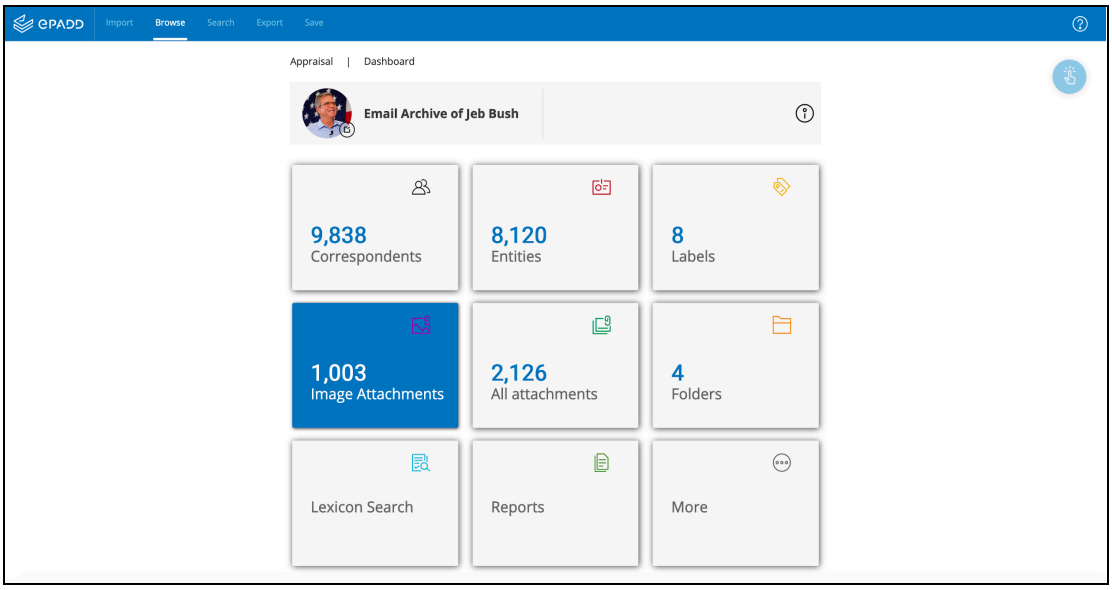
3.2.7.1 Reviewing Image Attachments

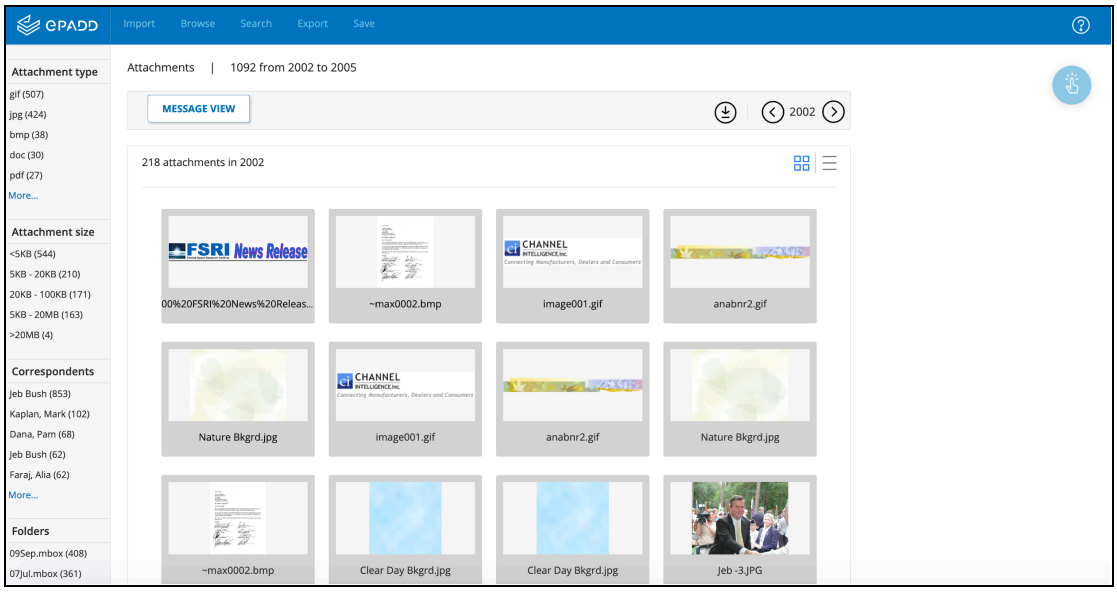
Select **Image Attachments** from the Appraisal Module menu to navigate to the Image Attachments screen. Image attachments are displayed chronologically by year and are de-duplicated. The user can navigate between years using the left and right arrows at the top right of the screen. The user can switch between a grid and list view using the icons in the top right corner. Selecting the Message View button will open the set of all messages associated with the currently displayed set of attachments. Selecting an image attachment will give the user the option to directly access the related message. The user can also export the current set of attachments as a .zip file by selecting the download icon.

Facets on the left can be used to further refine the set of images displayed by certain criteria including attachment type, attachment size, correspondent, and folder. Note that facets are tied to the messages associated with the currently displayed attachments. If the user selects '.pdf' from the attachment type facet, the resulting set of image attachments will include all attachments associated with messages that contain a .pdf attachment. This distinction means that the user will see attachments with other file types.

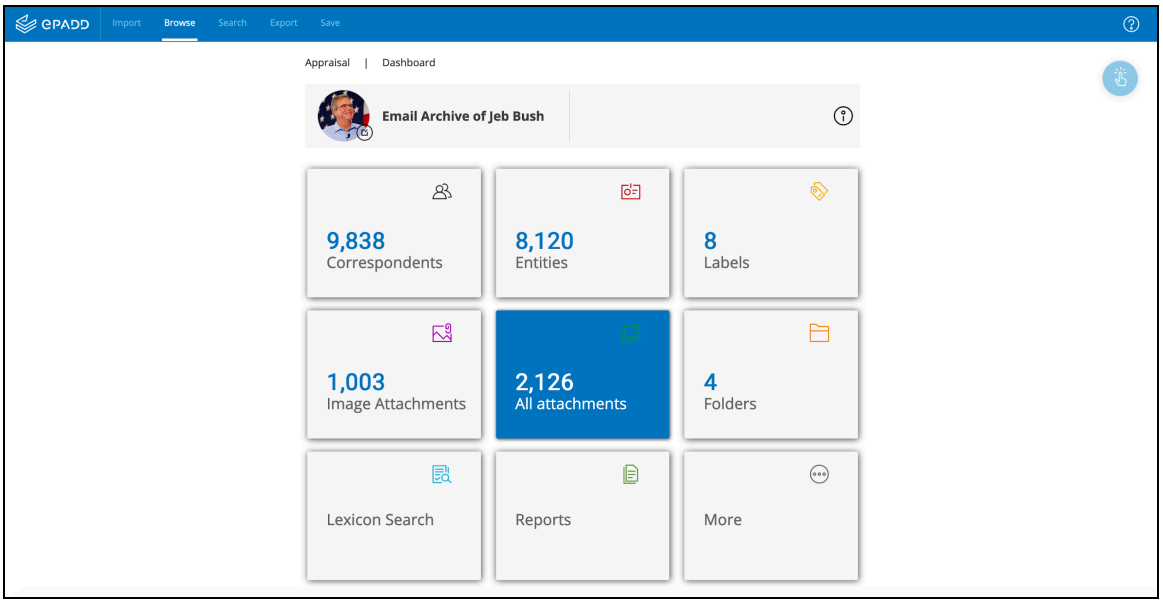
Pro-tip: You might use the file size and file type facets to exclude common non-controversial logos etc.

ePADD


ePADD
Import
Browse
Search
Export
Save
Appraisal | Dashboard
Email Archive of Jeb Bush
9,838
Correspondents
8,120
Entities
8
Labels
1,003
Image Attachments
2,126
All attachments
4
Folders
Lexicon Search
Reports
More


ePADD
Import
Browse
Search
Export
Save
Attachment type
gif (507)
jpg (424)
bmp (38)
doc (30)
pdf (27)
More...
Attachment size
<5KB (544)
5KB - 20KB (210)
20KB - 100KB (171)
5KB - 20MB (163)
>20MB (4)
Correspondents
Jeb Bush (853)
Kaplan, Mark (102)
Dana, Pam (68)
Jeb Bush (62)
Faraj, Alia (62)
More...
Folders
09Sep.mbox (408)
07Jul.mbox (361)
Attachments | 1092 from 2002 to 2005
MESSAGE VIEW
2002
218 attachments in 2002
00%20FSRI%20News%20Releas...
~max0002.bmp
image001.gif
anabnr2.gif
Nature Bkgrd.jpg
image001.gif
anabnr2.gif
Nature Bkgrd.jpg
~max0002.bmp
Clear Day Bkgrd.jpg
Clear Day Bkgrd.jpg
Jeb -3.JPG

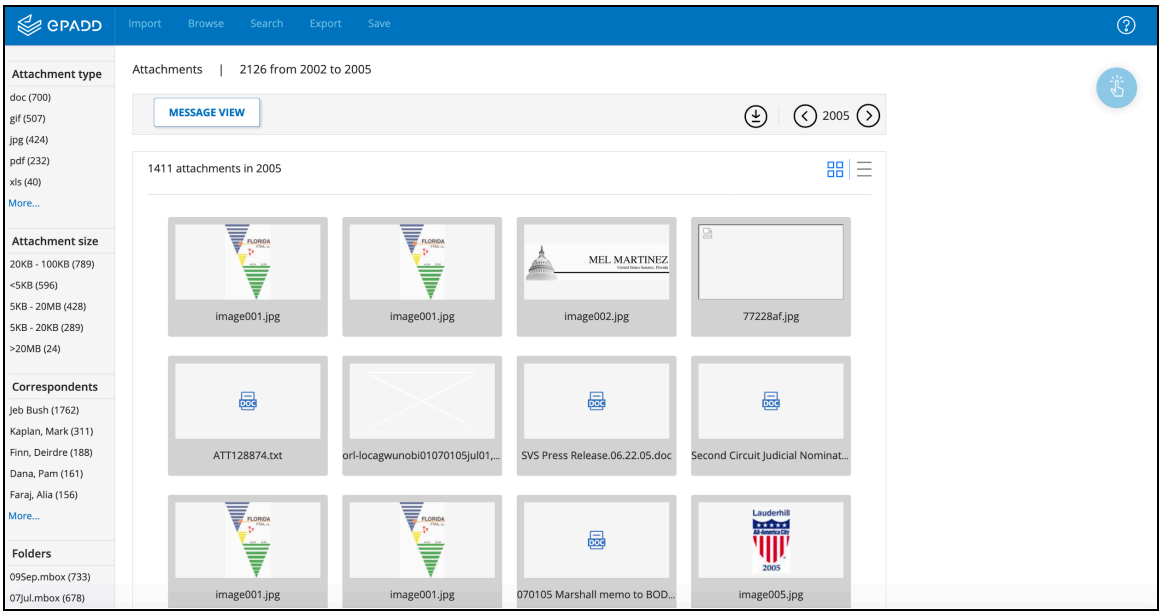
3.2.7.2 Reviewing All Attachments

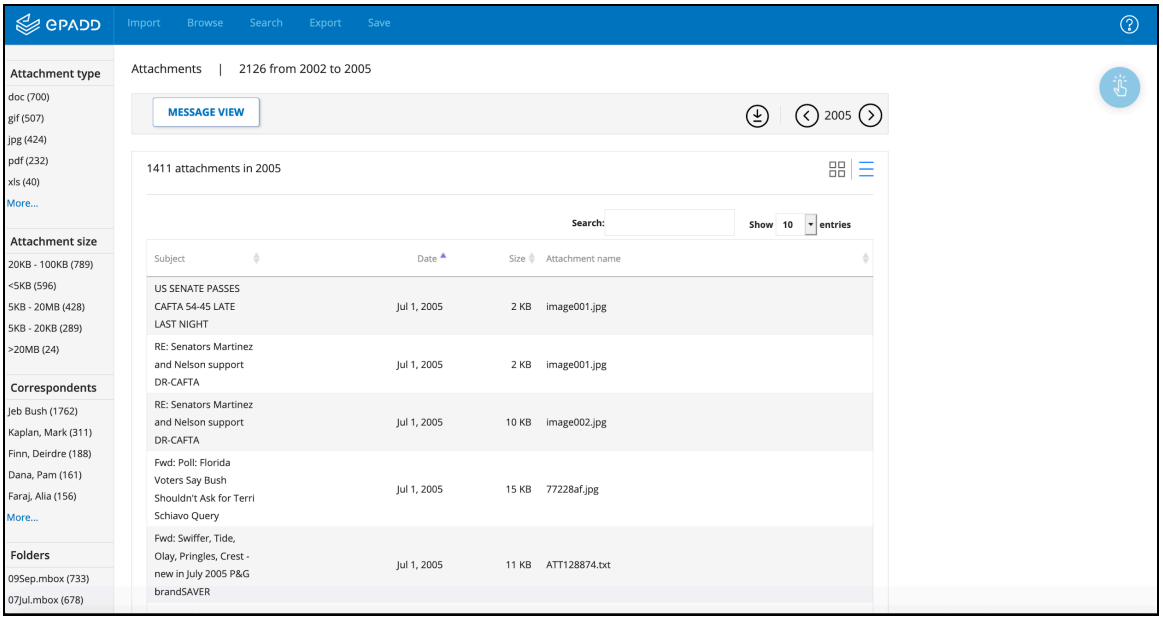


Select **All Attachments** from the Appraisal Module menu at the top of the screen to navigate to the All Attachments screen. This screen will present all attachments associated with the email archive, which can be viewed in a grid or list view, or downloaded and viewed outside of the ePADD program.

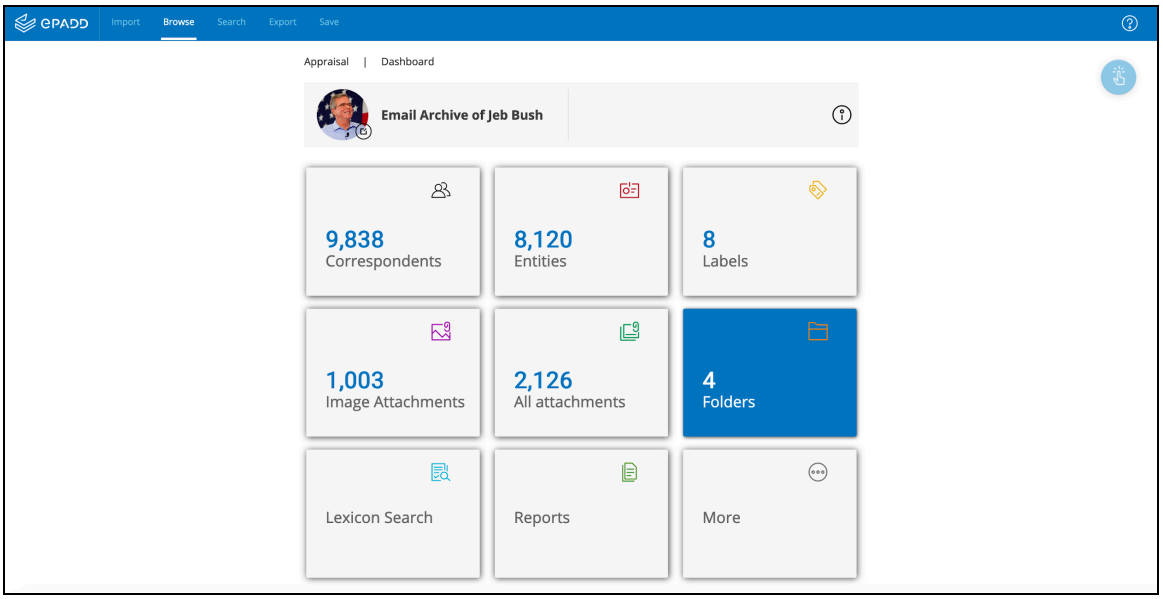
As with the Image Attachments view, attachments are displayed chronologically by year and are de-duplicated. The user can refine the set of attachments using the facets. The user can navigate between years using the left and right arrows at the top right of the screen. The user can switch between a grid and list view using the icons in the top right corner. Selecting the Message View button will open the set of all messages associated with the currently displayed set of attachments. Selecting an attachment will give the user the option to directly access the related message. Note that ePADD does not provide native support for viewing non-image attachments.

ePADD

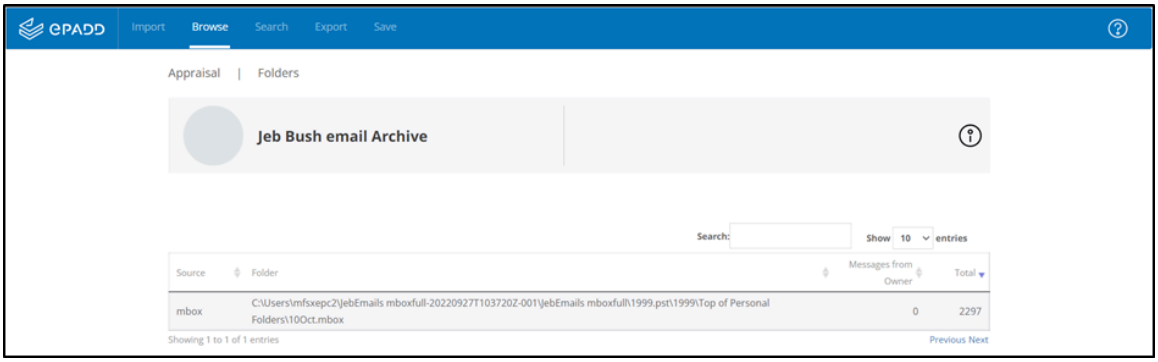

ePADD
Import
Browse
Search
Export
Save
Attachment type
doc (700)
gif (507)
jpg (424)
pdf (232)
xls (40)
More...
Attachment size
20KB - 100KB (789)
<5KB (596)
5KB - 20MB (428)
5KB - 20KB (289)
>20MB (24)
Correspondents
Jeb Bush (1762)
Kaplan, Mark (311)
Finn, Deirdre (188)
Dana, Pam (161)
Faraj, Alia (156)
More...
Folders
09Sep.mbox (733)
07Jul.mbox (678)
Attachments | 2126 from 2002 to 2005
MESSAGE VIEW
2005
1411 attachments in 2005
image001.jpg
image001.jpg
image002.jpg
77228af.jpg
ATT128874.txt
orl-locagwunobi01070105jul01,...
SVS Press Release.06.22.05.doc
Second Circuit Judicial Nominat...
image001.jpg
image001.jpg
070105 Marshall memo to BOD...
image005.jpg


ePADD
Import
Browse
Search
Export
Save
Attachment type
doc (700)
gif (507)
jpg (424)
pdf (232)
xls (40)
More...
Attachment size
20KB - 100KB (789)
<5KB (596)
5KB - 20MB (428)
5KB - 20KB (289)
>20MB (24)
Correspondents
Jeb Bush (1762)
Kaplan, Mark (311)
Finn, Deirdre (188)
Dana, Pam (161)
Faraj, Alia (156)
More...
Folders
09Sep.mbox (733)
07Jul.mbox (678)
Attachments | 2126 from 2002 to 2005
MESSAGE VIEW
2005
1411 attachments in 2005
Search:
Show 10 entries
Subject
Date
Size
Attachment name
US SENATE PASSES CAFTA 54-45 LATE LAST NIGHT
Jul 1, 2005
2 KB
image001.jpg
RE: Senators Martinez and Nelson support DR-CAFTA
Jul 1, 2005
2 KB
image001.jpg
RE: Senators Martinez and Nelson support DR-CAFTA
Jul 1, 2005
10 KB
image002.jpg
Fwd: Poll: Florida Voters Say Bush Shouldn't Ask for Terri Schiavo Query
Jul 1, 2005
15 KB
77228af.jpg
Fwd: Swiffer, Tide, Olay, Pringles, Crest - new in July 2005 P&G brandSAVER
Jul 1, 2005
11 KB
ATT128874.txt

3.2.8 Browsing Email Folders



The user can navigate email messages by folder structure using the Browse Folders interface, accessible from the Browse screen.



Browse Folders interface

If using Gmail, to ensure messages appear in the appropriate folder, please be sure to deselect the non-unique 'All Mail' folder during ingest for email ingested via IMAP.

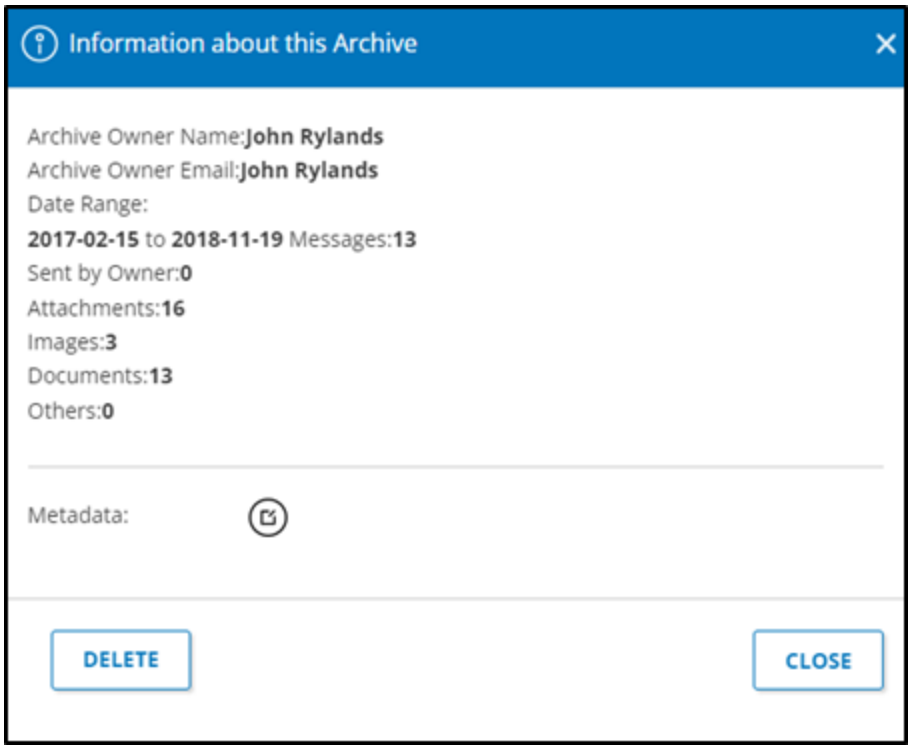
3.3 Appraising Email

Once you have confirmed the identity of correspondents, edited the lexicon to your satisfaction, and reviewed entities and attachments, the next step is using the information you have learned to appraise individual email messages. ePADD provides the user with a faceted search interface from which to annotate individual email messages and make appraisal decisions.

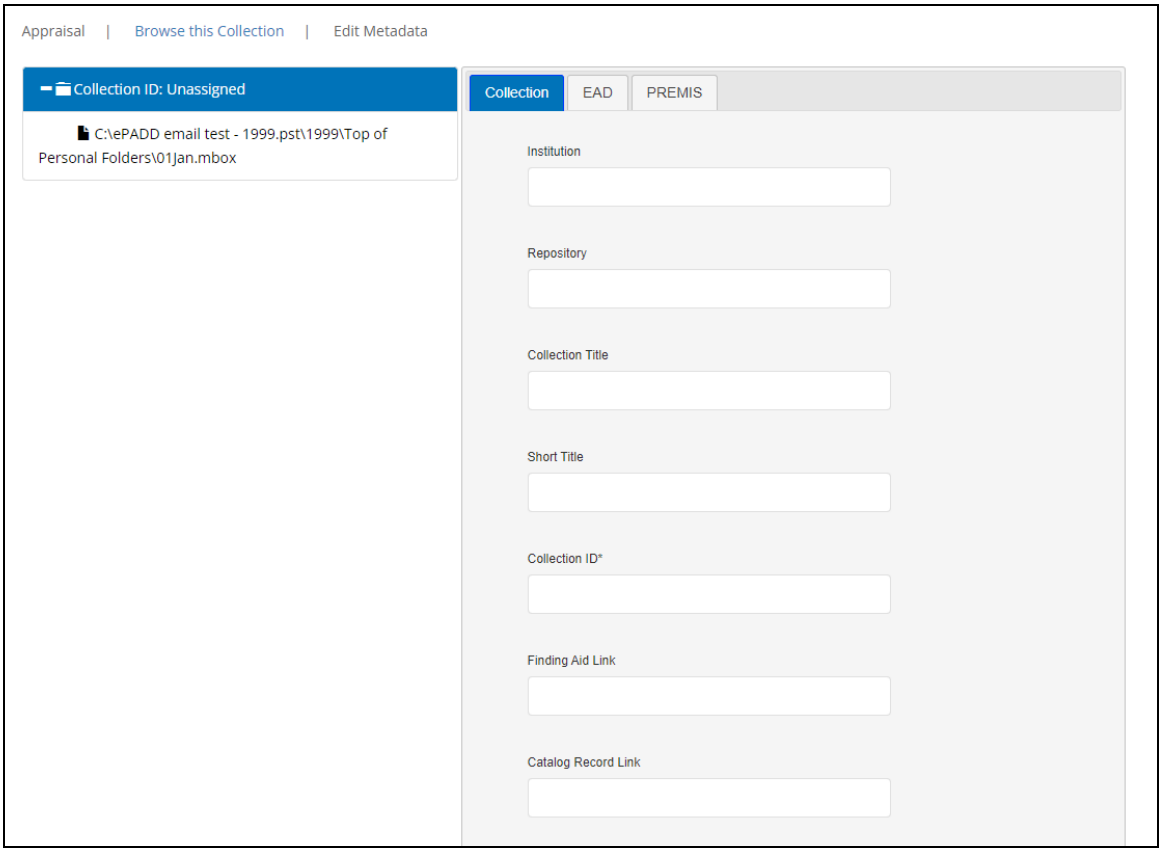
3.3.1 Reviewing and Creating Collection Metadata

ePADD allows users to add manually an extensive range of metadata associated with a particular email archive. First, click on the information button - - which is located to the right of the email archive's title on the ePADD dashboard:

This will open a dialog box entitled 'Information about this Archive'. Here, click on the edit button - - found next to the heading 'Metadata':

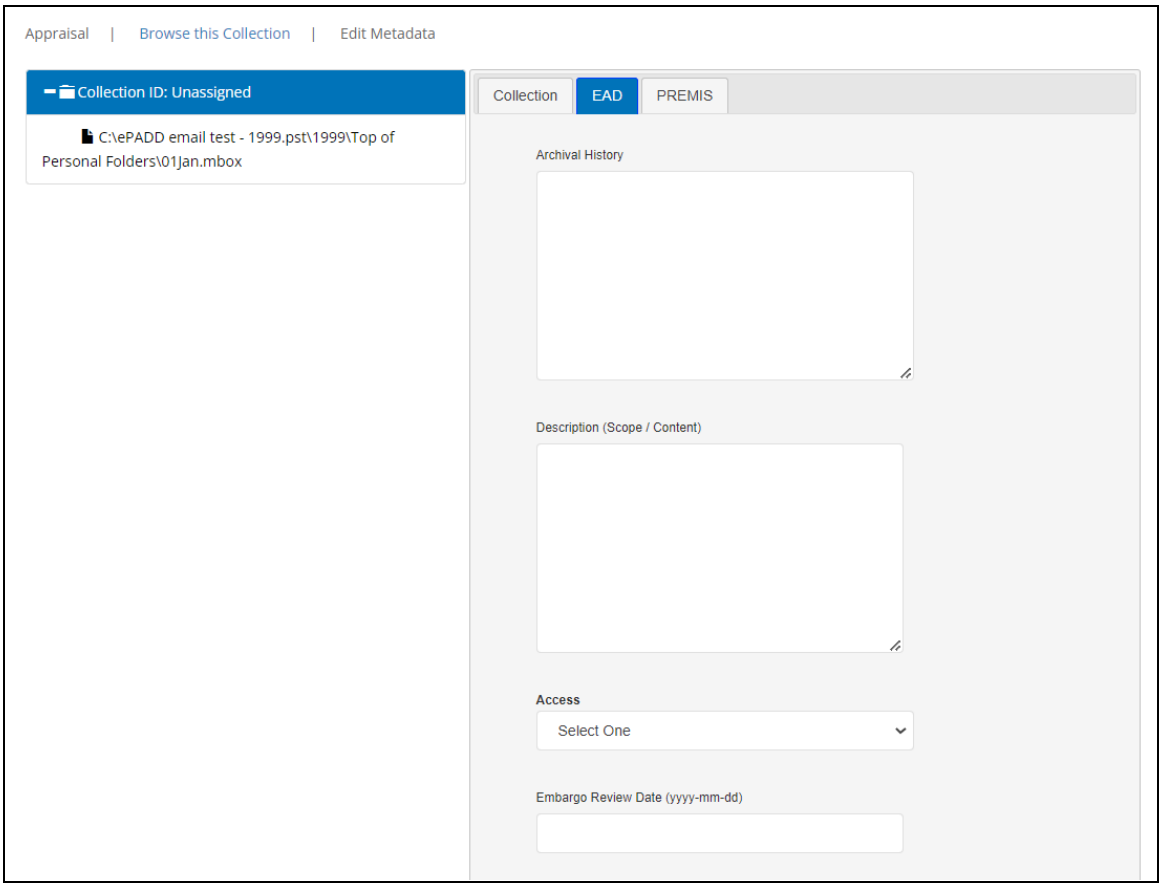


This presents you with a variety of metadata options:

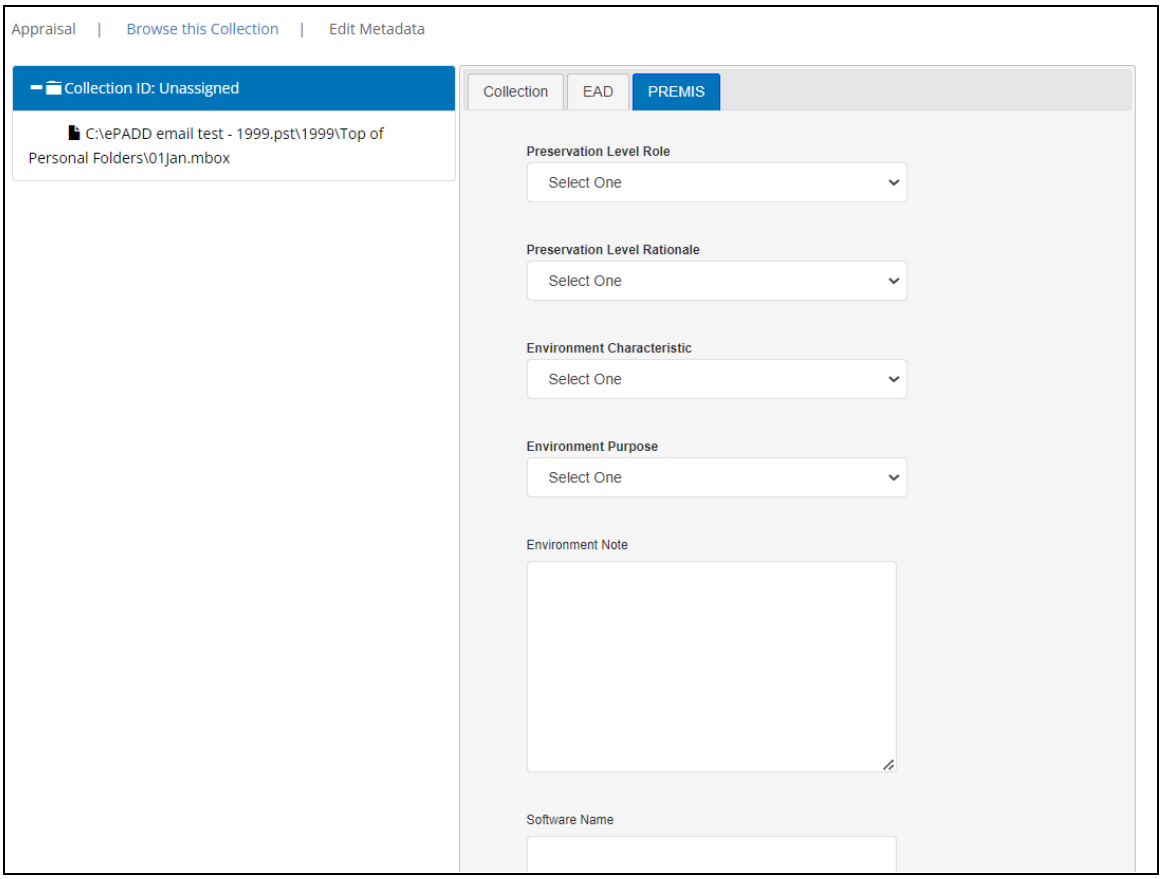


On the tabs you can switch between Collection, EAD (Encoded Archival Description) and PREMIS. In the Collection metadata you can add some basic information about the collection, such as Collection Title, short description, and scope and content, as well provide links to online catalog descriptions.

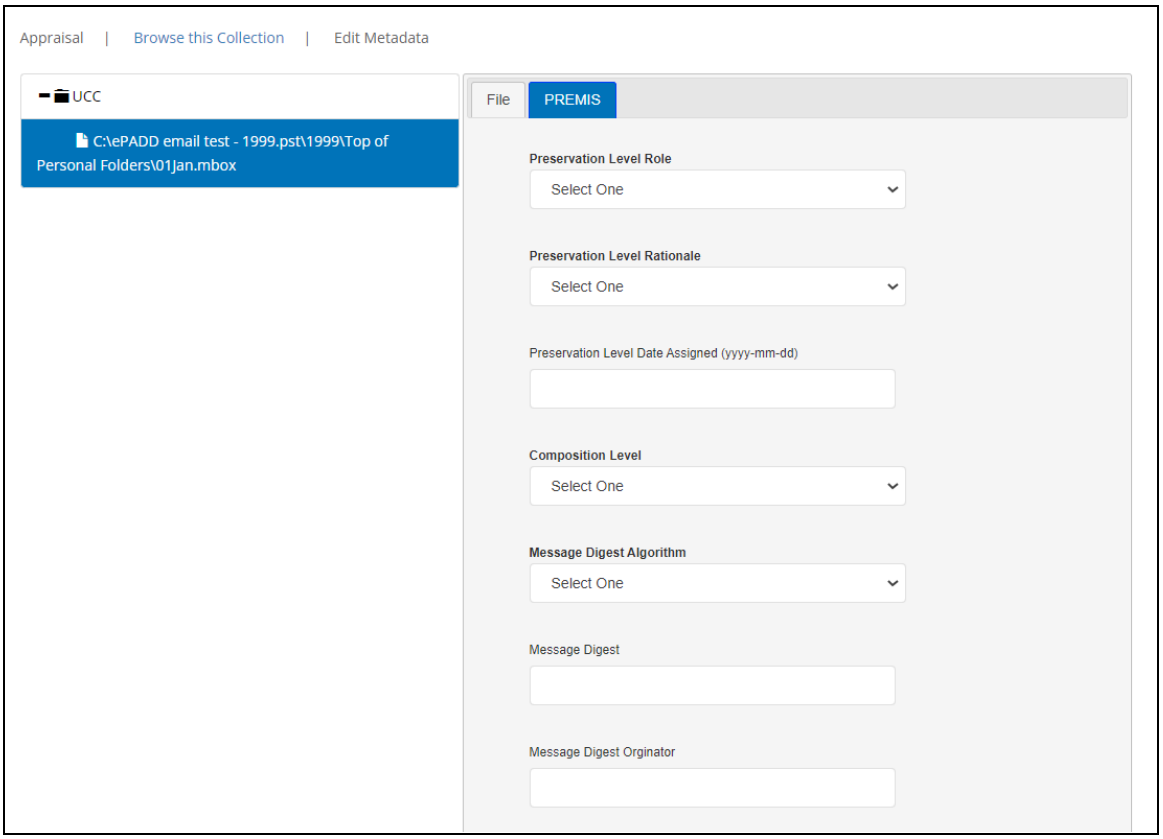
You can do the same with metadata elements associated with EAD by selecting the EAD tab:


Appraisal | Browse this Collection | Edit Metadata
Collection ID: Unassigned
C:\ePADD email test - 1999.pst\1999\Top of Personal Folders\01Jan.mbox
Collection
EAD
PREMIS
Archival History
Description (Scope / Content)
Access
Select One
Embargo Review Date (yyyy-mm-dd)

And with PREMIS:



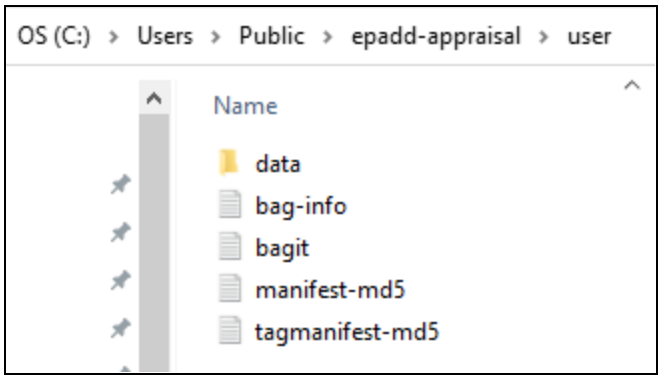
You can add or edit PREMIS metadata for an individual file, not just the whole collection, by selecting one of the files listed to the left of the metadata fields. This will present you with two tabs, 'File,' where you can add notes about a particular file, and 'PREMIS,' where you can add metadata relevant to that particular file:



PREMIS also has an XML schema that has been developed for preservation metadata. The xml file generated by ePADD is found in the folder 'user' in the directory:

On Windows: C:\Users\<username>\epadd-<module>\
On Mac: /Users/<username>/epadd-<module>/

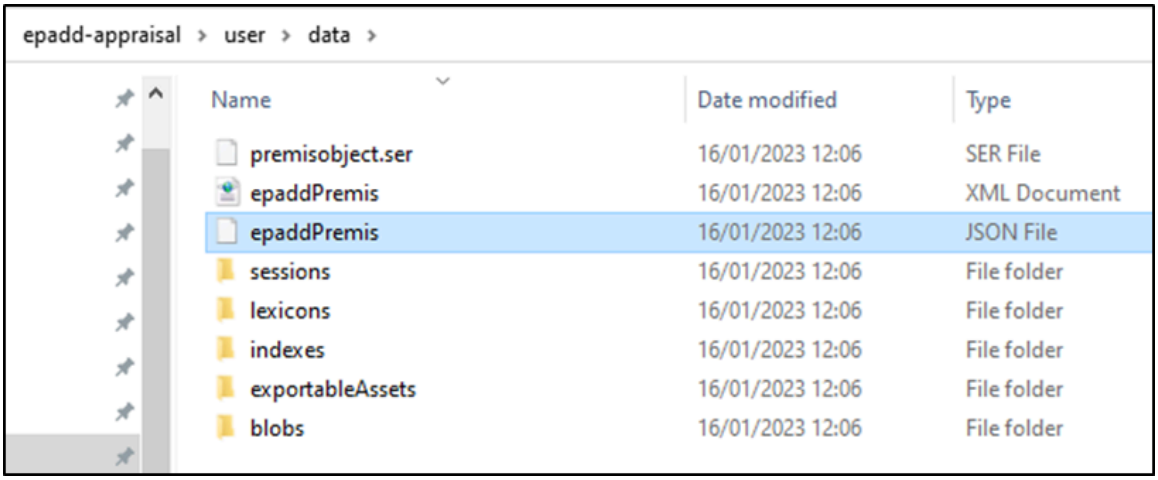
The file epaddPremis.xml is located in In the folder 'data':



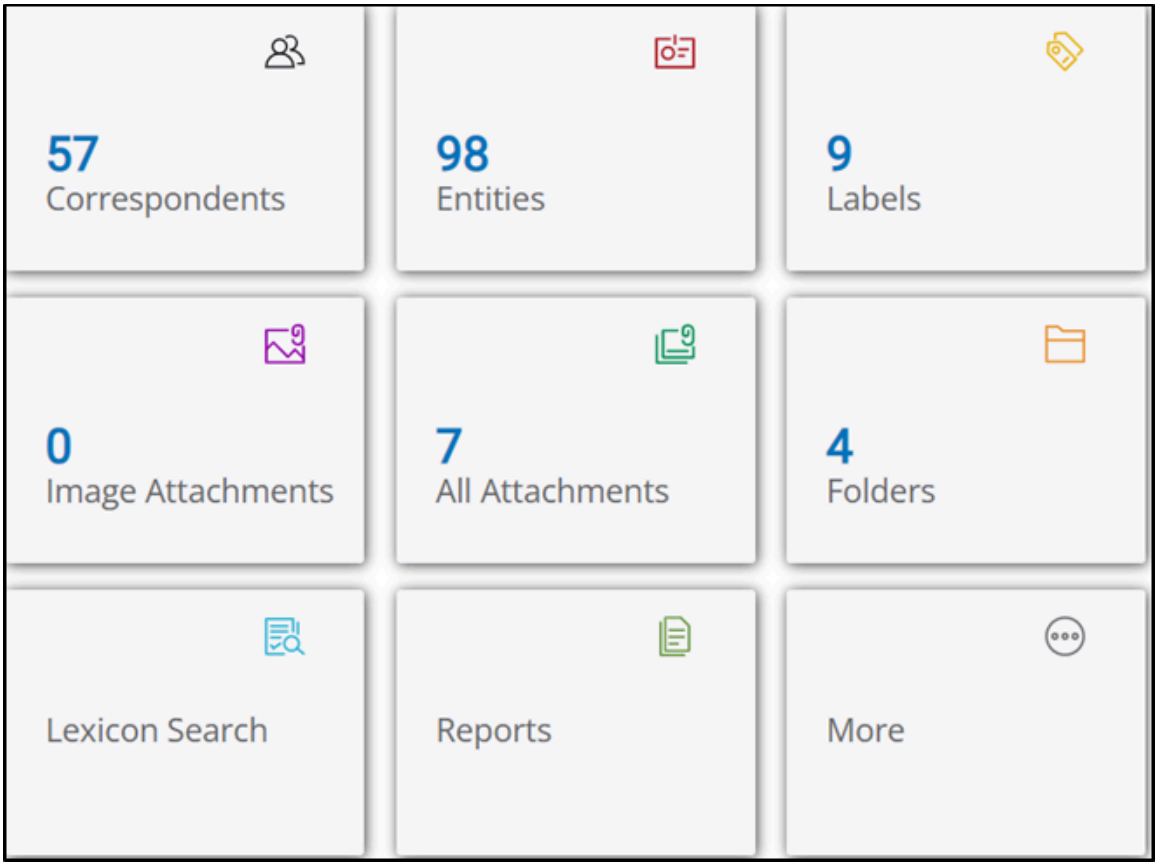
Here you can see an example of the xml file, which contains the manually added PREMIS metadata as well as some metadata automatically added by ePADD, such as date and time of ingest, number of attachments in the collection:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<premis:premis xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org/1999/
https://www.loc.gov/standards/premis/premis.xsd" version="3.0">
    <premis:object xsi:type="premis:intellectualEntity">
        <premis:objectIdentifier>
            <premis:objectIdentifierType>local</premis:objectIdentifierType>
            <premis:objectIdentifierValue></premis:objectIdentifierValue>
        </premis:objectIdentifier>
        <premis:preservationLevel>
            <premis:preservationLevelValue>bit level</premis:preservationLevelValue>
            <premis:preservationLevelRole authority="preservationLevelRole" authorityURI="http://id.loc.
            "http://id.loc.gov/vocabulary/preservation/preservationLevelRole/reg"></premis:preservationL
            <premis:preservationLevelRationale></premis:preservationLevelRationale>
            <premis:preservationLevelDateAssigned>2022-06-20</premis:preservationLevelDateAssigned>
        </premis:preservationLevel>
        <premis:significantProperties>
            <premis:significantPropertiesType>image/jpg</premis:significantPropertiesType>
            <premis:significantPropertiesValue>1</premis:significantPropertiesValue>
        </premis:significantProperties>
        <premis:significantProperties>
            <premis:significantPropertiesType>application/zip</premis:significantPropertiesType>
            <premis:significantPropertiesValue>2</premis:significantPropertiesValue>
        </premis:significantProperties>
        <premis:significantProperties>
            <premis:significantPropertiesType>application/msexcel</premis:significantPropertiesType>
            <premis:significantPropertiesValue>1</premis:significantPropertiesValue>
        </premis:significantProperties>
```

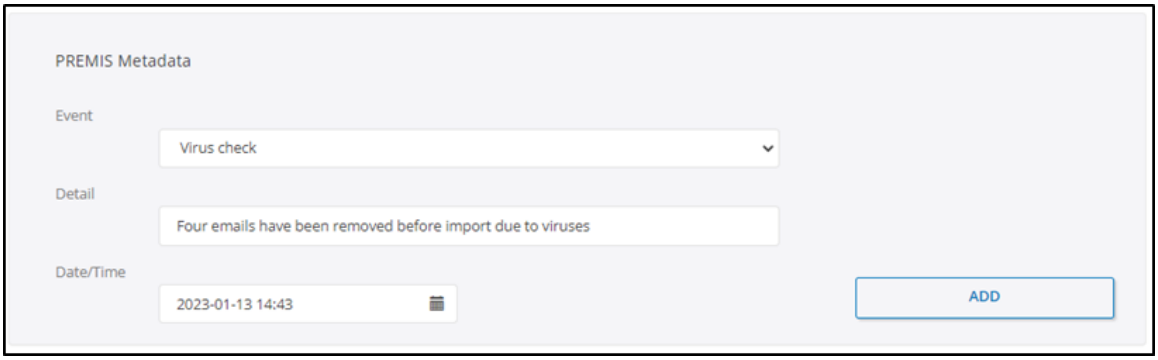
As well XML, PREMIS metadata is also available in JSON format.



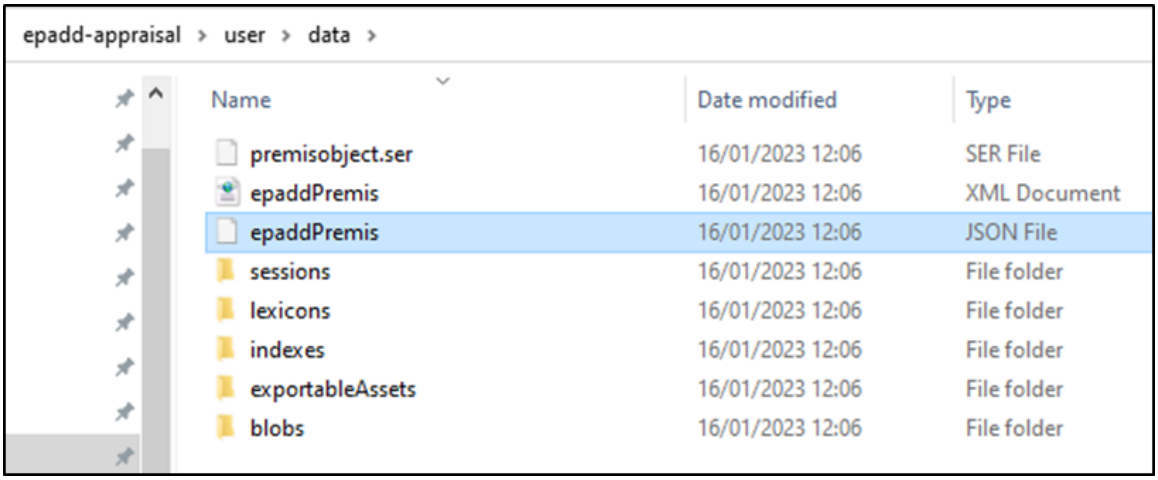
To add a PREMIS event manually, select 'More' on the ePADD dashboard:



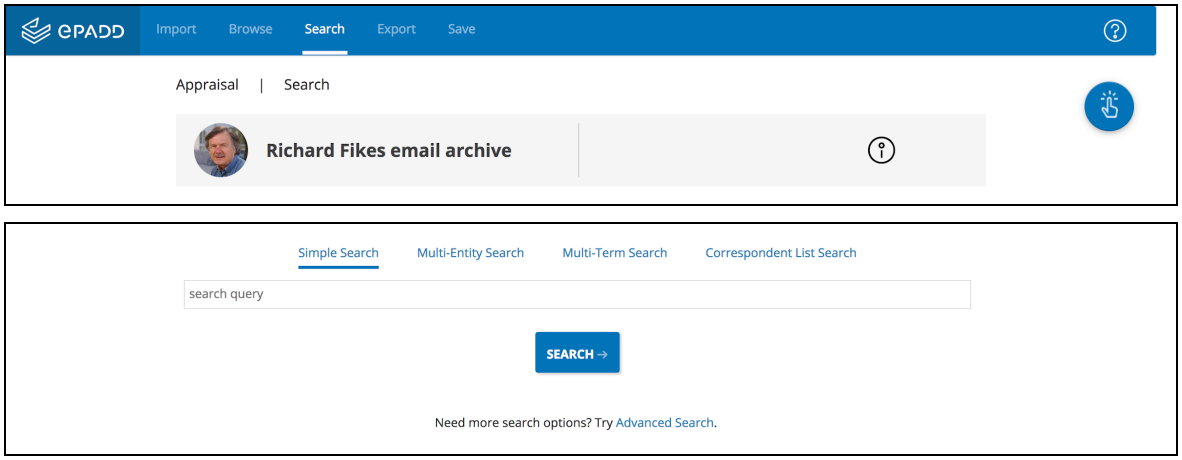
Here, users can select a PREMIS event type, such as 'Virus check,' from the dropdown list, add further information to the free text 'Details' field, and enter the date/time of the event.



As well XML, PREMIS metadata is also available in JSON format.



3.3.2 Searching Email Messages



To search the email archive, navigate to the Search screen from the Appraisal Module Menu.

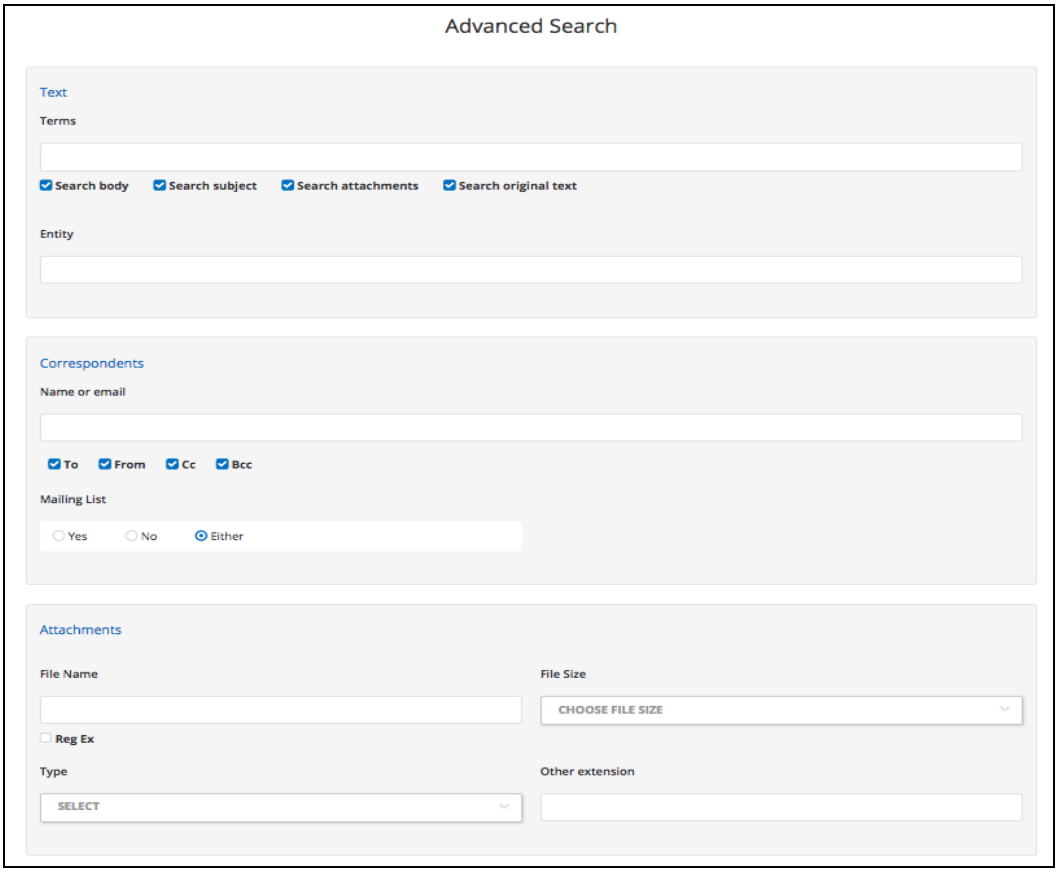
Besides lexicons, accessible via the Browse page, ePADD supports five types of search: simple search, advanced search, multi-entity search, multi-term search, and correspondent list search.

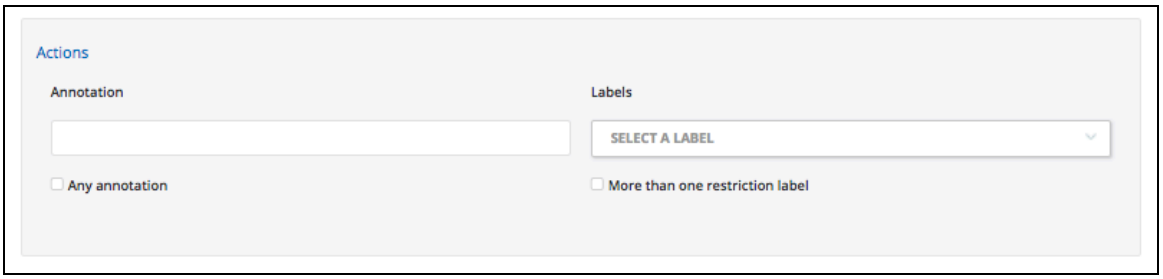
To use the simple search feature, simply type in your query into the simple search box available on the initial search screen. This will search across message headers (correspondents and subjects) and message text, as well as the text of all document attachments.

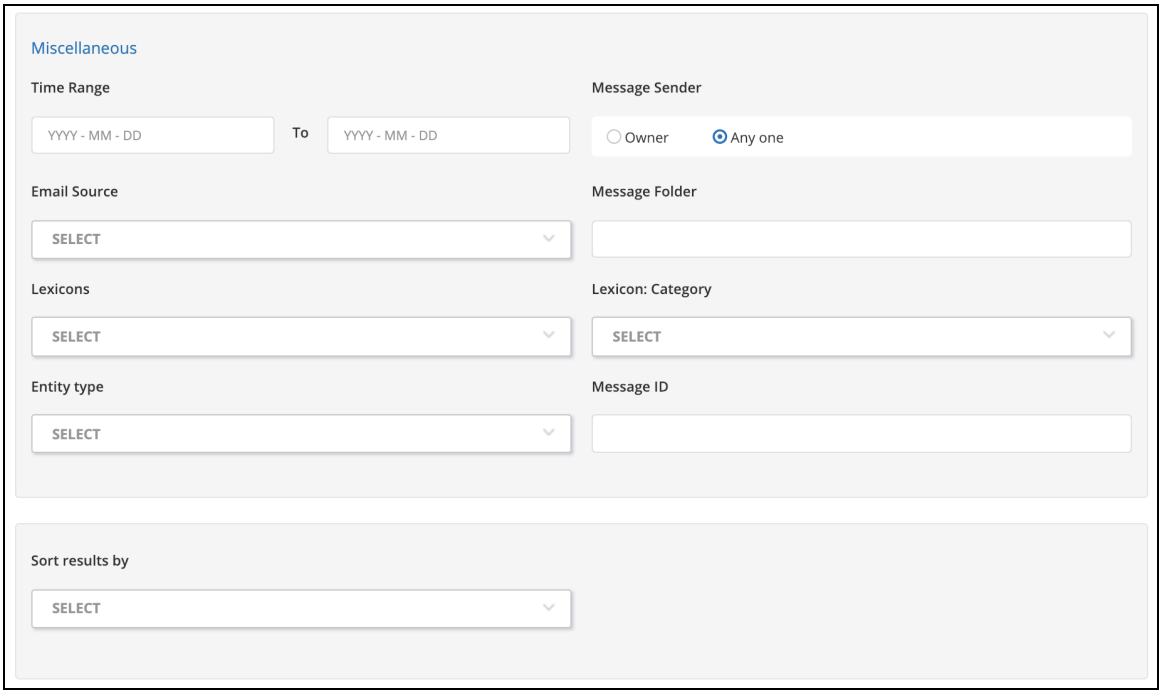
Once you have entered your term(s) and made your selection, click the *Search* button. The search string will appear highlighted on the message browse screen.

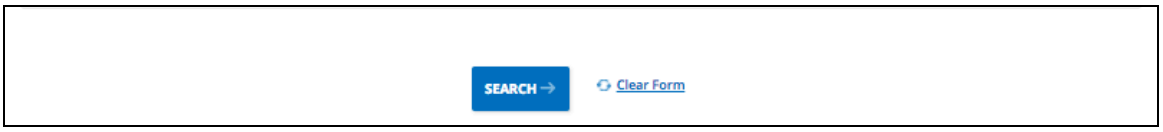
Advanced Search enables the user to generate complex searches of email text,

correspondents, attachments, and any actions previously taken on the messages, as well as a variety of other parameters.









Search Tips can be accessed by selecting the help button in the top right corner of the *Advanced Search* screen.

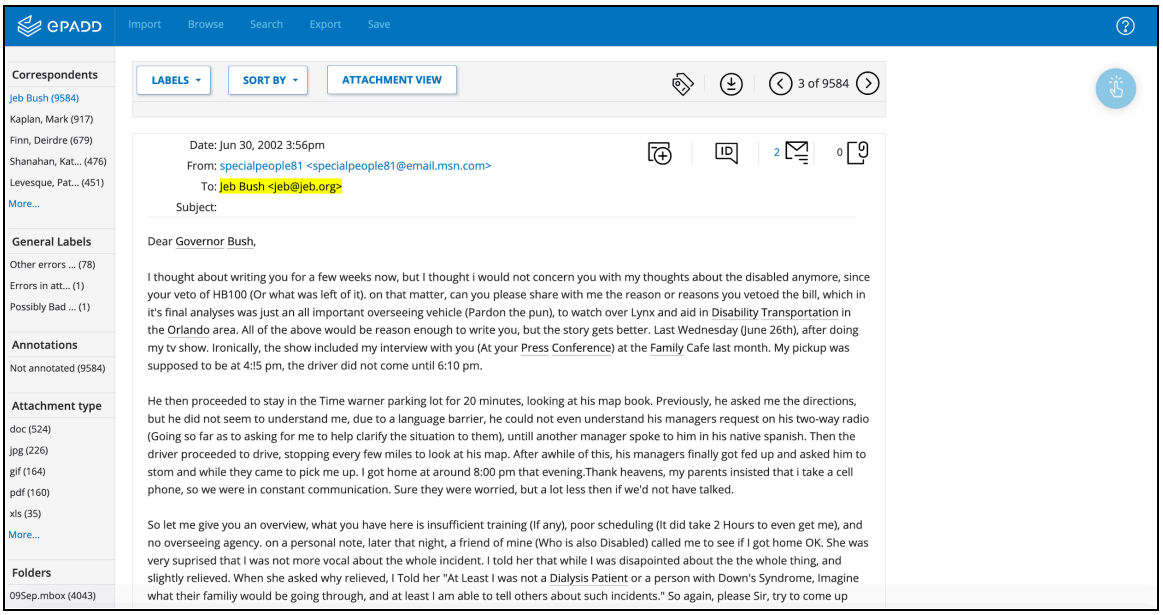
ePADD also allows a user to search the email archive for entities contained in a given block of provided text, using *Multi-entity Search*. Any matching terms will be highlighted on the results page, and hyperlinked to corresponding messages.

Likewise, ePADD enables a user to search an email archive for multiple terms using *Multi-term Search*. Type or paste one term per line. Any matching terms will be highlighted on the results page, and hyperlinked to corresponding messages.

Finally, ePADD enables a user to search for a list of email addresses using Correspondent List Search. Type or paste one term per line to search for and apply labels to a set of email addresses.

3.3.2.1 Navigating Search Results

Upon submitting a search query, ePADD opens a new browser window displaying the message browse screen.

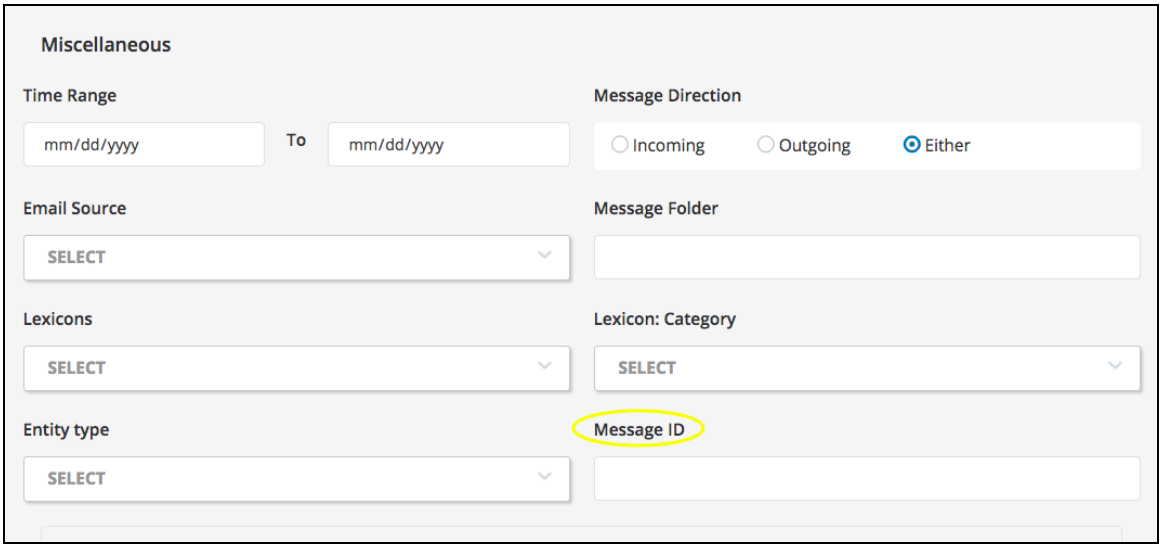


The message browse screen displays the full text of email messages as well as a faceted browser which allows the user to further limit the search results. The total number of messages resulting from your search appears above the email message display. The current message appears as a fraction of that display. For instance, if you are appraising message 1 out of 400 results, the display would read 1/400. The forwards and back arrows to the left and right of that number allow you to navigate between search results.

Search results can be limited by correspondent, email direction (incoming or outgoing), lexicon category (if relevant), attachment type (if relevant), annotations or the originating mail folder. To limit the results by a particular facet, simply click the term from the left-hand panel.

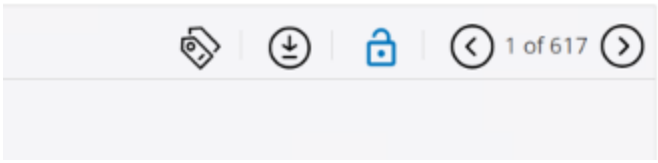
All messages have been assigned a unique ID, accessible from this page. Select the ID button to copy the ID to your clipboard:

If you ever wish to return to a particular message, you can search for a message ID (or semicolon delimited set of message IDs) using the Advanced Search interface in the Miscellaneous search category.

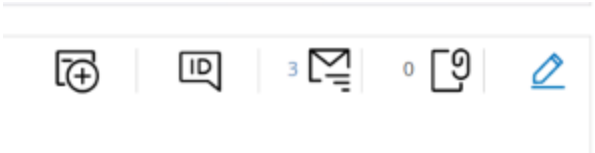


3.3.3 Redacting Email Messages

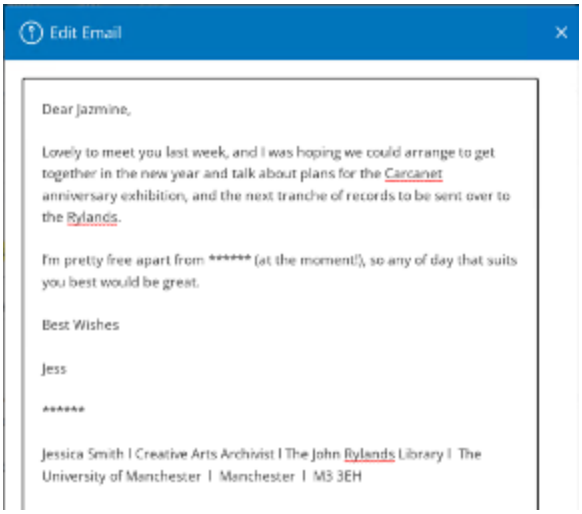
Email redaction is available in the processing module. ePADD will save a 'preservation' copy and a 'normal' copy of an email message, to ensure that message integrity is preserved. There are also an option to export the preservation or normal versions of the email from ePADD as MBOX files, depending on which version is preferred by the organization.
Ensure that the padlock icon is set to open, which indicates the message is in normal view:



Select the 'modify email content' icon



This produces the following screen, where text of the message can be edited and redacted. When editing is complete, select 'apply' and a redacted copy of the message will be saved.



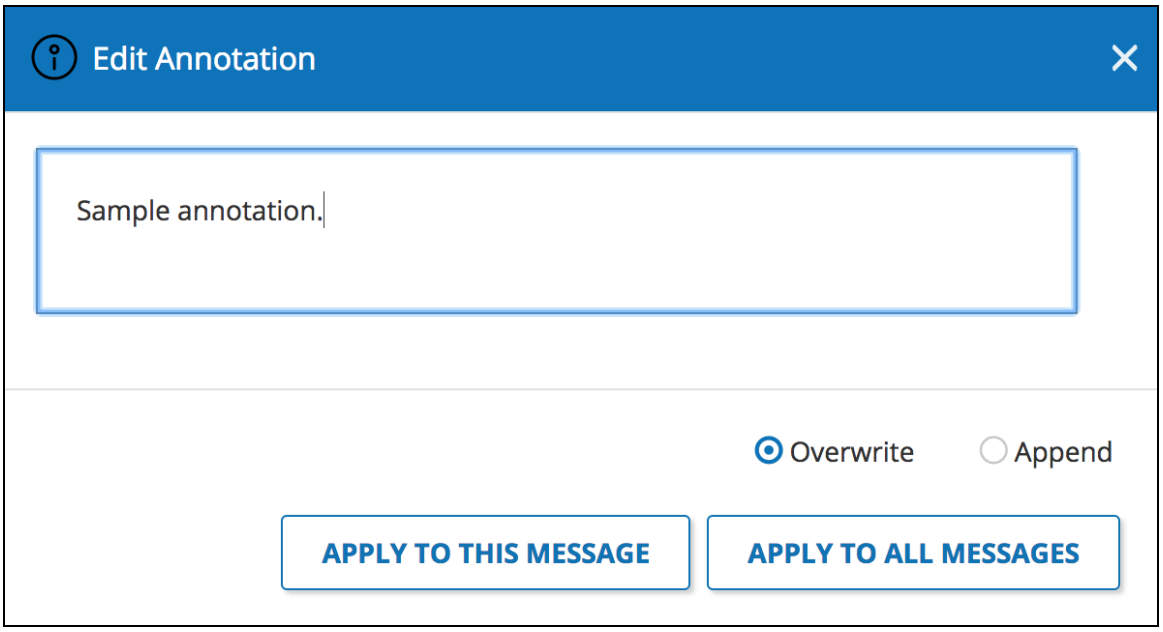
Viewing redacted messages

By selecting the padlock icon, users can toggle between the preservation and normal versions of a message, and select the view that they prefer for their purposes.

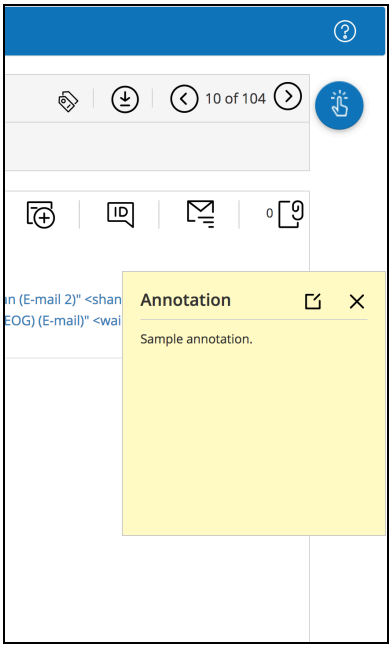
3.3.4 Annotating Email Messages

ePADD allows the user to annotate email messages from the message browse screen. These annotations are persistent across ePADD sessions, and can be used to provide supplemental information to assist the archivist in processing the collection, including elaborating on access restrictions or to provide additional information for the benefit of researchers.

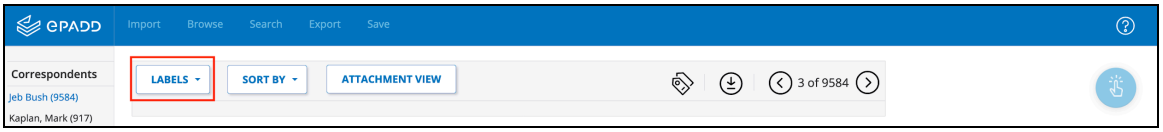
To annotate an email message, click the annotation button in the bar above the message headers: . Users can select to apply the annotation to a single message or all messages currently being reviewed. Likewise the user can either append the annotation to current annotations, or overwrite all existing annotations with the new annotation. **Warning: Overwriting annotations cannot be undone!**



Once you have added an annotation, a new facet, Annotations, will appear in results on the message browse screen. Additionally, when you navigate to that message or those messages in the future, the annotation will be visible.



3.3.5 Labeling Email Messages

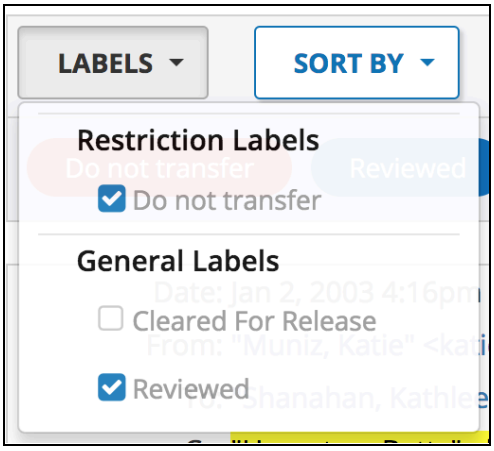


ePADD allows the user to label email messages that should not be exported for transfer to an archival repository, as well as messages that should be transferred with restrictions.

To label a message, navigate to the message or set of messages, then select a label or set of labels from the Labels dropdown.

There are three default labels within ePADD: Do Not Transfer is a default, not editable restriction label that indicates the message should not be transferred to the next module(s). Cleared for Release is a default, not editable restriction label that indicates the message that was previously labeled with at least one restriction is now cleared for release and should be transferred to the next module(s).

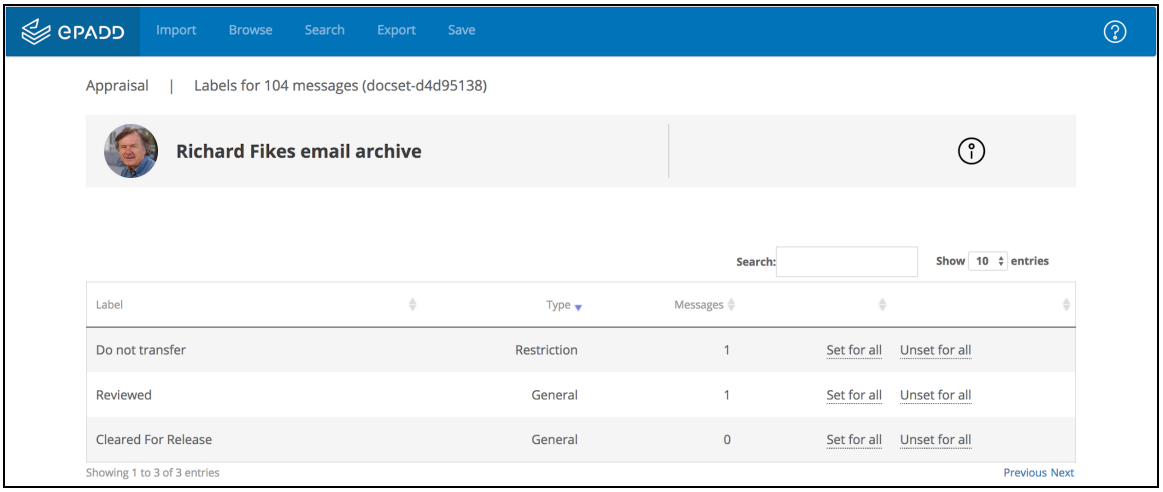
Reviewed is a default, editable label that indicates the message being browsed has been reviewed, and can be used to help track one's progress in reviewing the email archive.



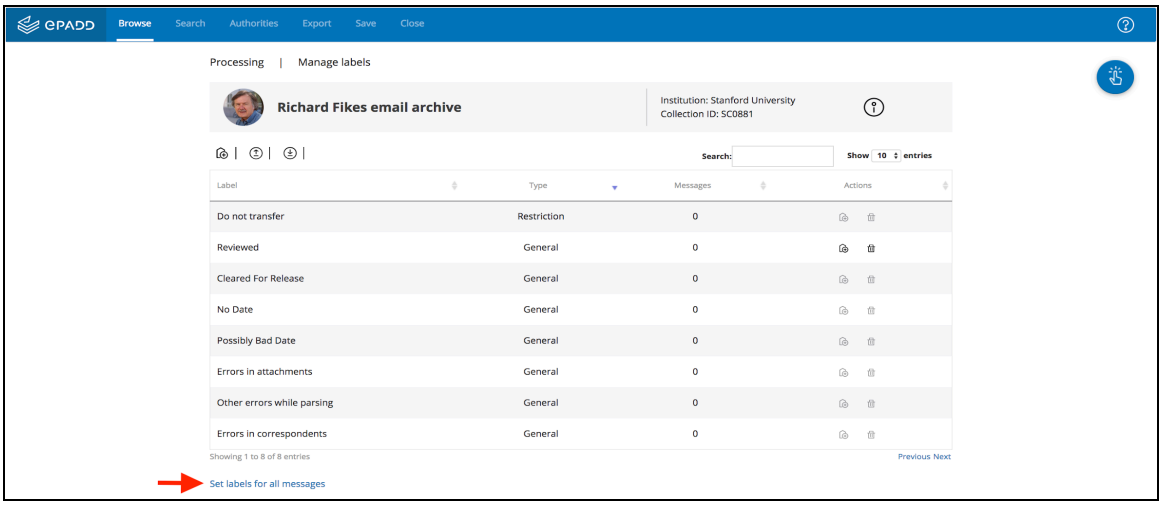
After you have assigned labels, they appear in the message browse screen.

3.3.5.1 Assigning bulk labels

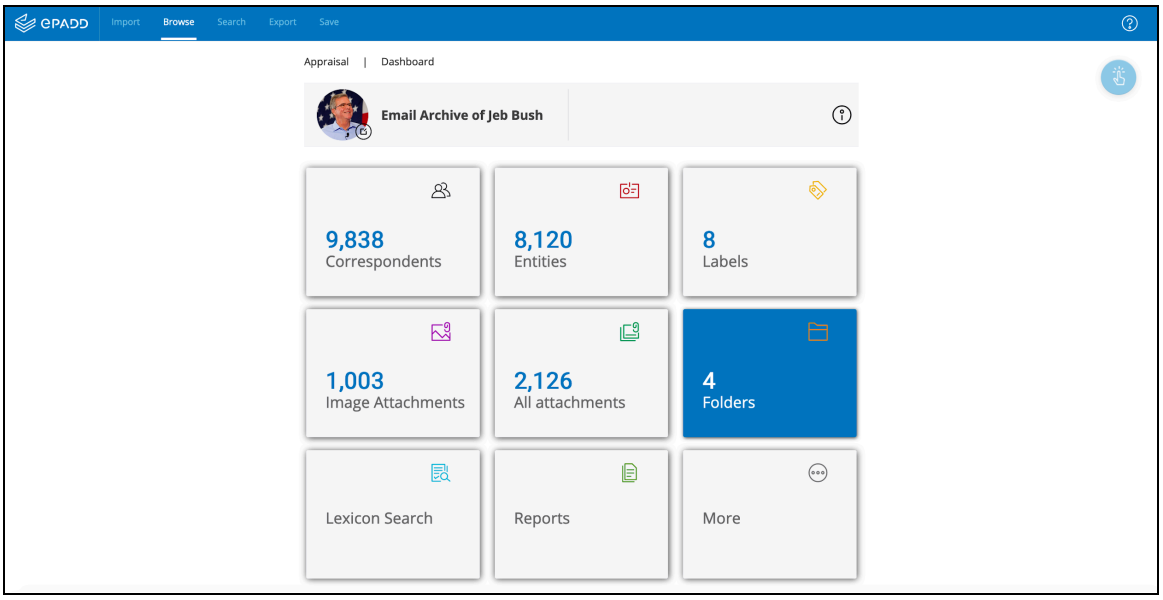
To apply a label to more than one message, you can select the label management button above the message pane:

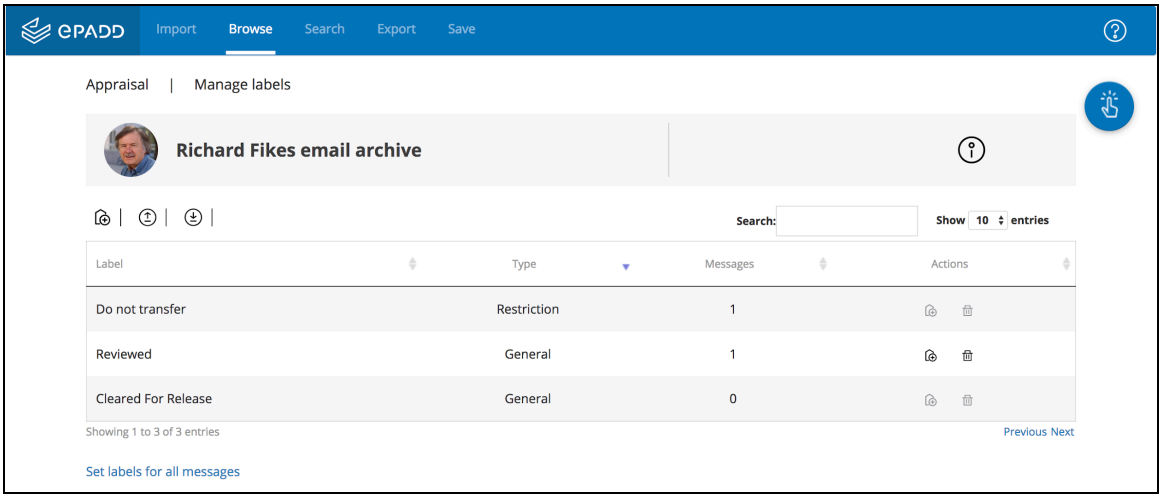


You can also assign default labels for all messages by using the 'Set labels for all messages' function.



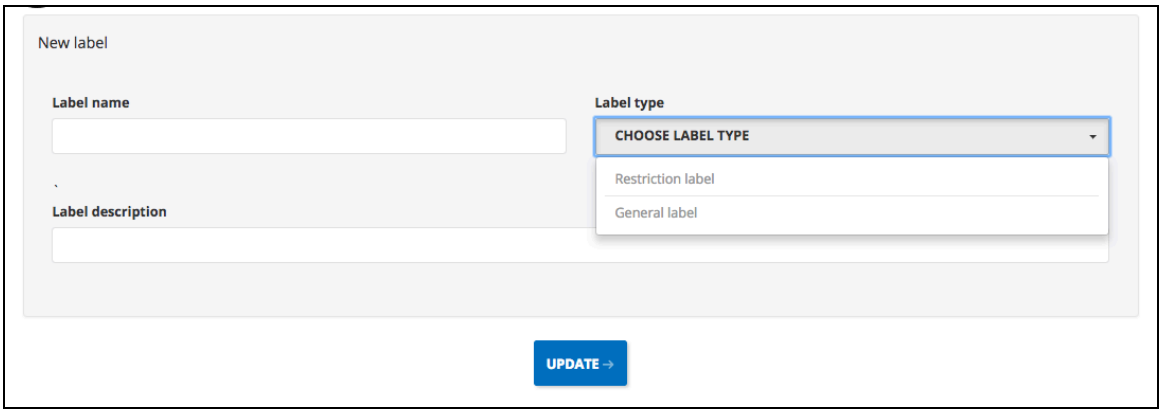
3.3.5.2 Creating a new label

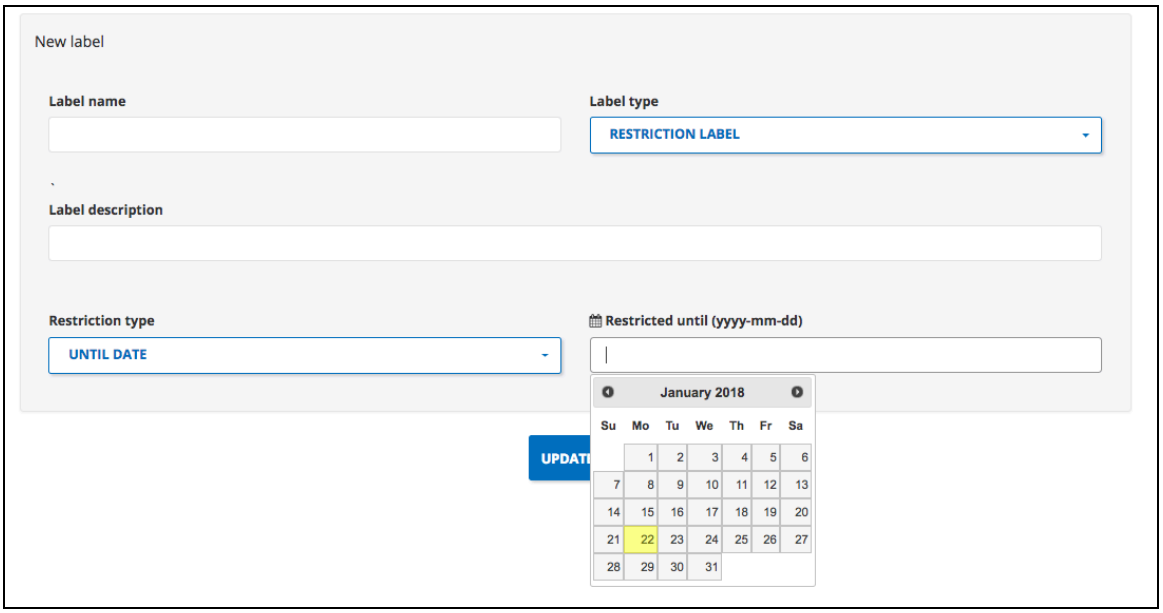




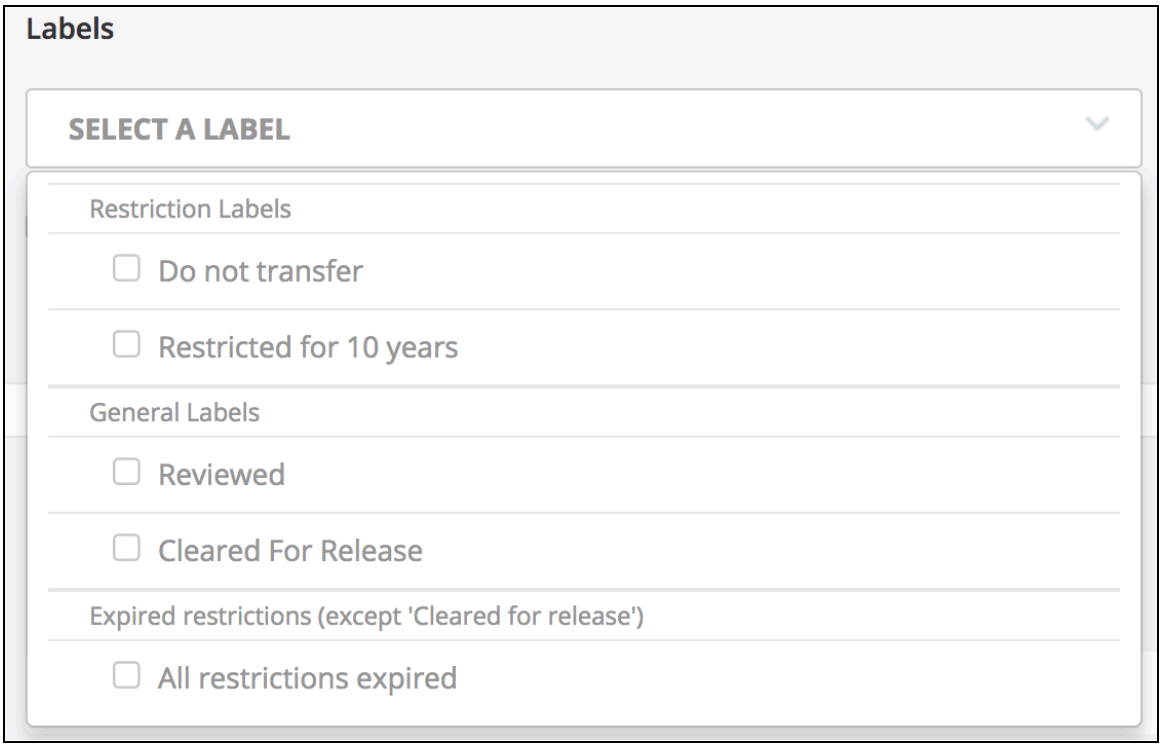
Users can create new labels, import labels (overwriting all current label assignments) or export all labels via the Labels interface.

Below are several screenshots illustrating how to create a new restriction label.



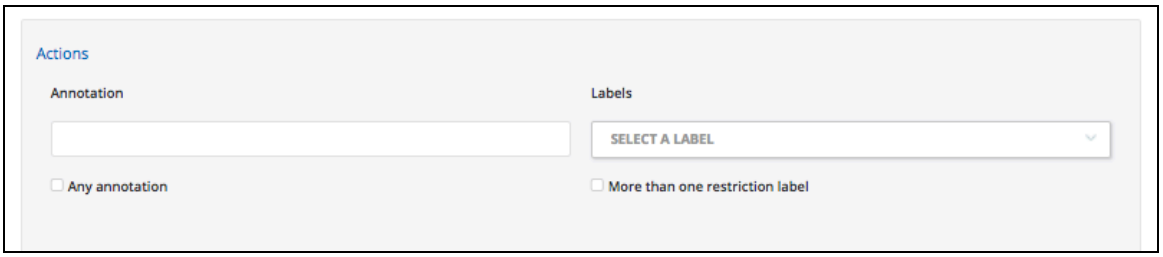


Messages that are restricted until a particular date, or restricted for a given time period from the date of message creation, are considered machine-actionable, as ePADD is able to recognize when the restriction period has expired via an advanced search function.



Messages marked with restriction labels will not automatically export to Discovery and Delivery. In order to export these messages, the user needs to manually assign them the label 'Cleared for Release.' Note that ePADD does not delete any messages that have restriction labels applied to them, they are simply not exported to subsequent modules.

Because ePADD enables users to assign multiple restriction labels to a given message or set of messages, you may find it helpful to review the set of messages that contains more than one restriction. You can perform this search query under Advanced Search.



Viewing all messages with more than one restriction via advanced search.

Warning: Since importing labels overwrites all current label assignments, this task should only be undertaken prior to accessioning / processing a collection.

Known issue: The introduction of message error labels in v7 has complicated the label upload mechanism. For the time being, please refrain from using the Upload Label Description feature on the Manage Labels screen. You can manually add or delete labels as needed as a workaround.

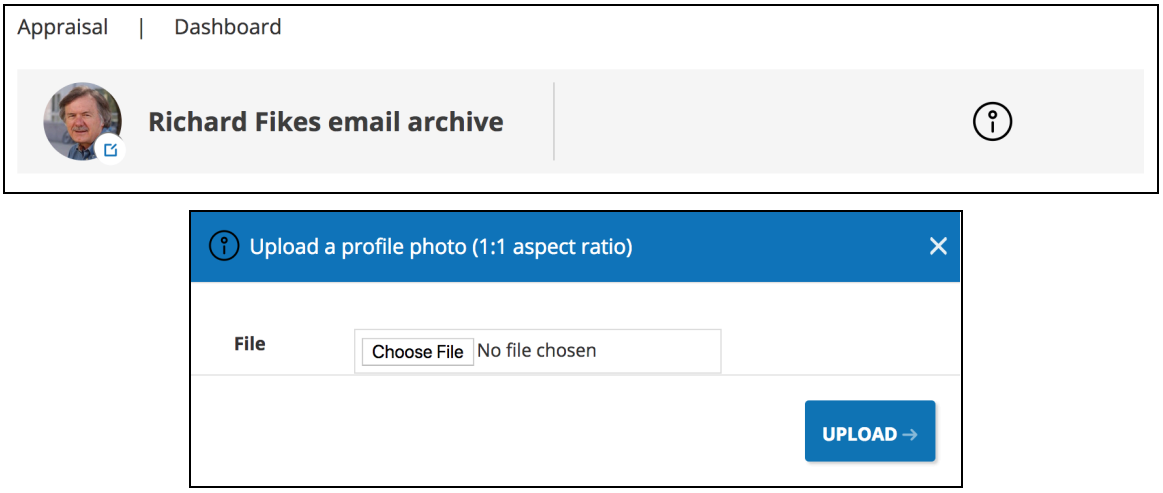
3.3.5.3 Assigning bulk labels by correspondent

ePADD also gives the user the ability to assign labels to all messages from a list of correspondents, using the *Correspondents List Search* option.

3.3.6 Assigning Images to the Collection

ePADD allows the user to assign up to three images to any given collection to improve the end-user experience: a profile image in Appraisal, and profile, landing page, and banner images in Processing. Images must be resized to the dimensions indicated to display correctly.

Images are set in the location they appear. You can set the profile image from the browse page using the edit option within the greyed out profile image.

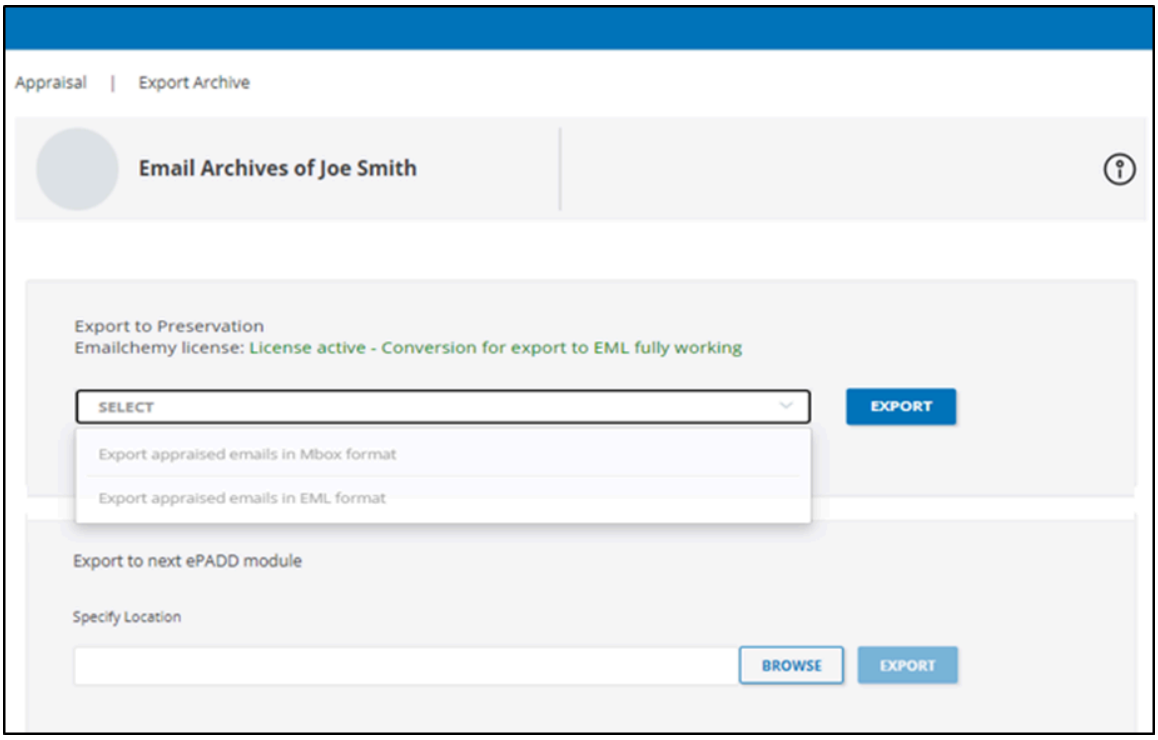


Within Processing, you can add the landing image from the Collection Browse page, and the banner image from the Collection Details page, using a similar interface.

3.4 Exporting Email for Preservation

ePADD can be used to create and export a preservation package or 'preservation ready bag', complete with all email headers, that can be transferred to a digital preservation repository. An email archive can be exported as an MBOX file or, if the Emailchemy license is active, as an EML file. To create the preservation package click **Export** on the ePADD dashboard:

In 'Export to Preservation', select either 'Export appraised emails in MBOX format' or 'Export appraised emails in EML format' and click 'Export':



When files are imported into the Appraisal module, ePADD creates what we refer to as a Canonical-Acquisitioned file. 'Canonical' is the single file format designated by ePADD as a common standard based format for transfer and long-term preservation purposes. The designated ePADD 'Canonical' format is MBOX. 'Acquisitioned', in this context, refers to the unmodified form of email as originally received by ePADD and not yet subject to any appraisal or processing. The folder:

On Windows: C:\Users\<username>\epadd-appraisal\user\data\exportableAssets
On Mac: User/<username>/epadd-appraisal/user/data/exportableAssets

contains the folders:

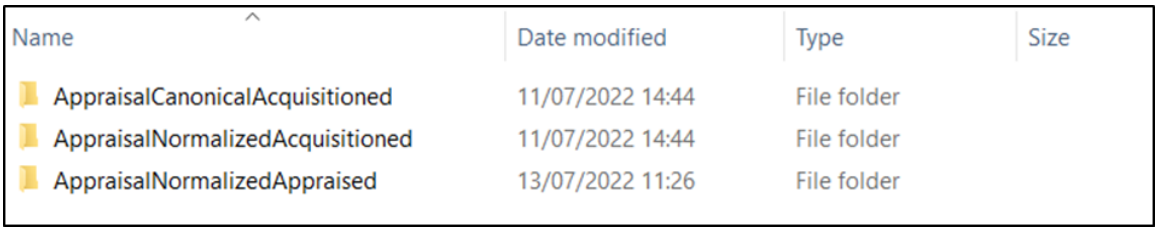
- AppraisalCanonicalAcquisitioned

- AppraisalNormalizedAcquisitioned

Clicking on 'Export' creates a folder – 'AppraisalNormalizedAppraised' – containing the appraised MBOX file(s), which is found via the following file path:

On Windows: C:\Users\<username>\epadd-appraisal\user\data\exportableAssets\AppraisalNormalizedAppraised

On Mac: User/<username>/epadd-appraisal/user/data/exportableAssets/AppraisalNormalizedAppraised



The AppraisalCanonicalAcquisitioned folder contains the folders:

- files before conversion with emailchemy
- untouched imported mbox files

These folders contain the original, unprocessed files. The folder AppraisalNormalizedAcquisitioned contains the folders:

- mbox files generated by emailchemy
- untouched imported mbox files

These folders contain files converted to MBOX by Emailchemy, as well as original, unprocessed MBOX files.

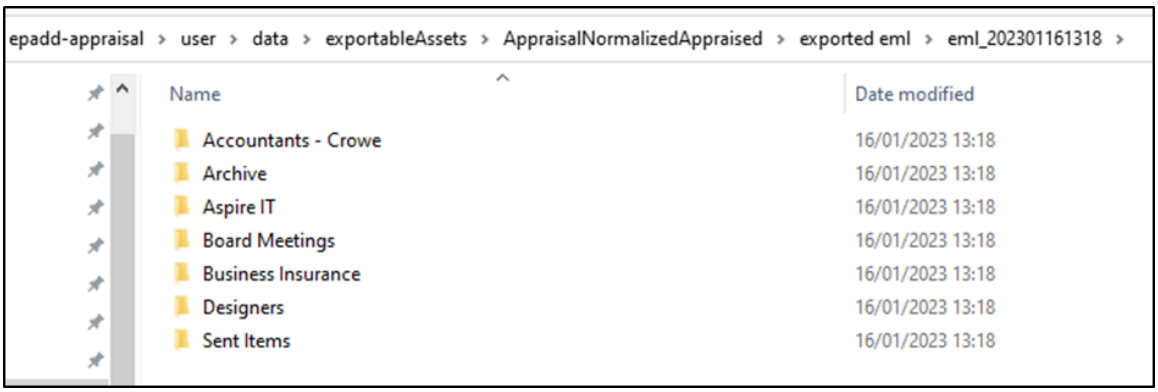
The folder 'AppraisalNormalizedAppraised' contains appraised emails in their respective folders and will include any attached labels or annotations.

When emails are exported in the EML format, ePADD creates a sub-directory in AppraisalNormalizedAppraised folder:

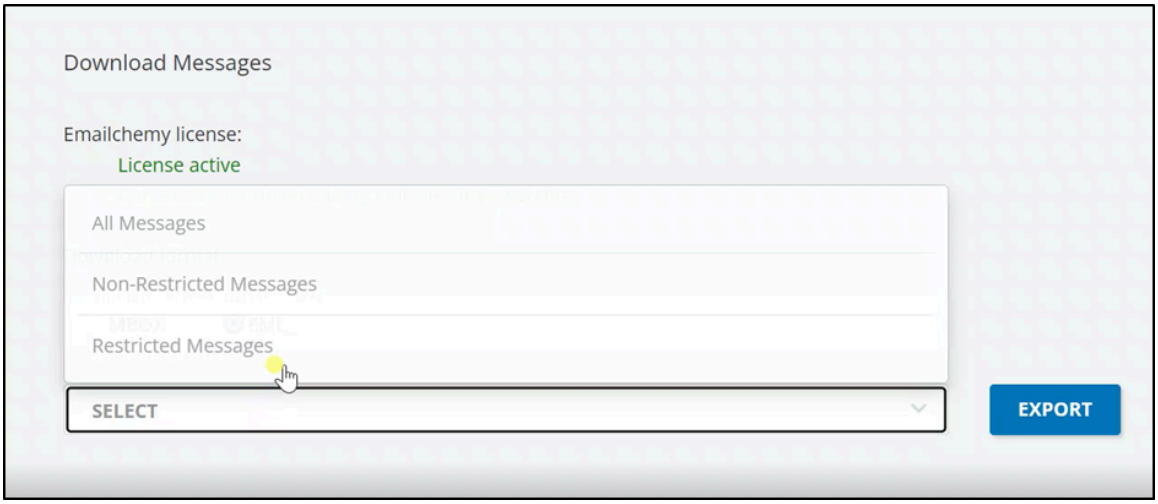
On Windows: C:\Users\<username>\epadd-appraisal\user\data\exportableAssets\AppraisalNormalizedAppraised\exported eml

On Mac:
User/<username>/epadd-appraisal/user/data/exportableAssets/AppraisalNormalizedAppraised/exported eml

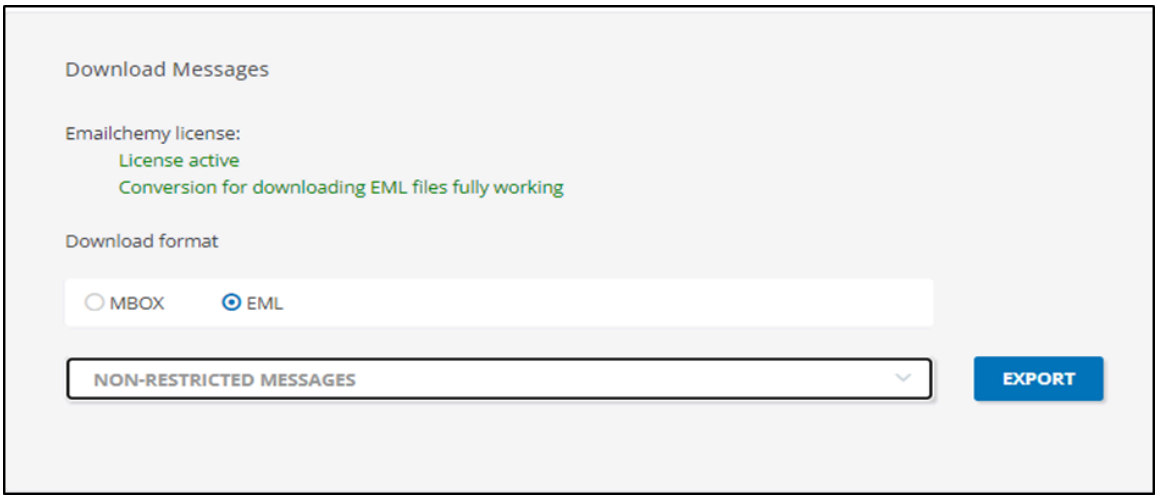
Here is an example of the folder structure within the EML folder:

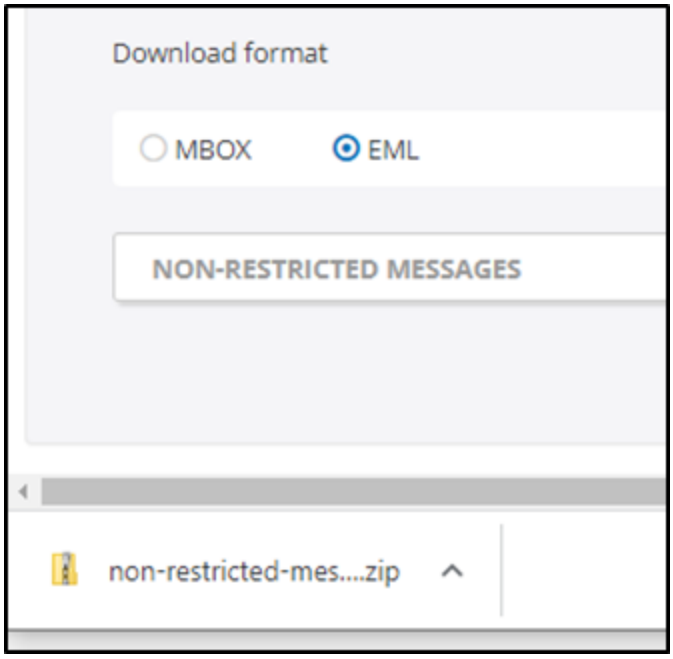


Emails can also be downloaded in MBOX or EML format via ePADD's Export page. If the MBOX format is selected, one MBOX file containing all of the emails is downloaded. An Emailchemy license is required if emails are to be downloaded in EML. Users can select 'All Messages', 'Non-Restricted Messages' or 'Restricted Messages':



If the EML format is selected, the emails are downloaded as a zip file:

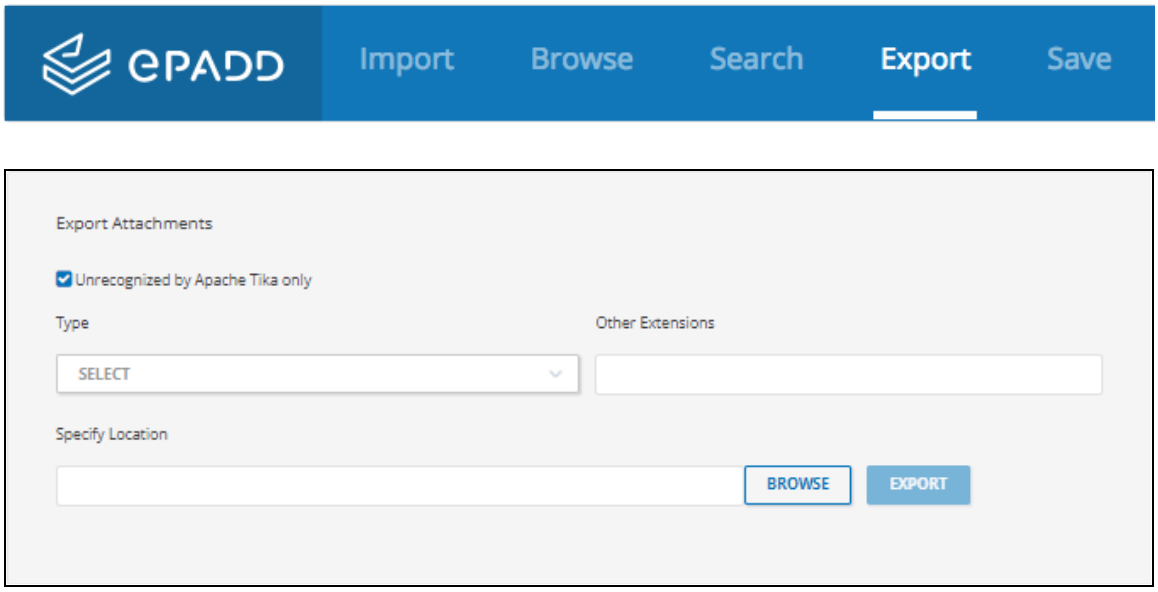




If MBOX is selected, then one Mbox file with all emails is downloaded.

3.5 Exporting Message Attachments

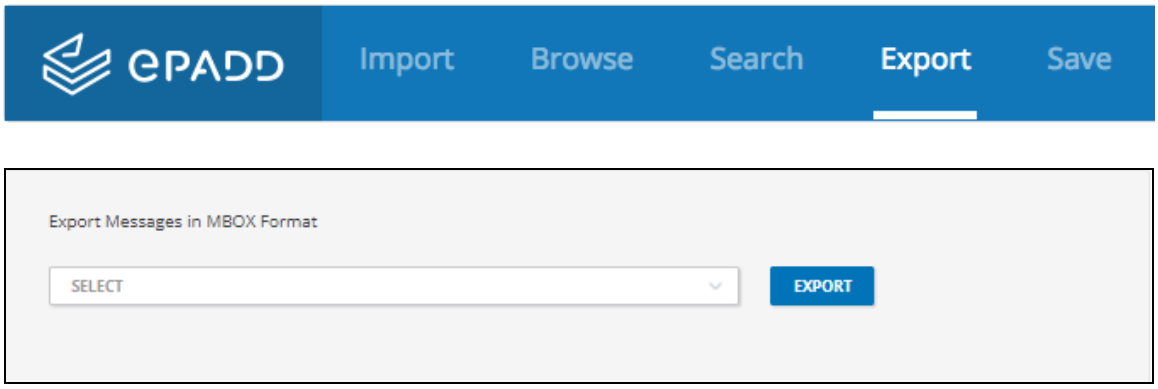
Select Export from the Appraisal Module menu at the top of the screen to navigate to the Export screen.



You can choose to export attachments according to various criteria, including particular formats. The user can also indicate whether they would like to export only those attachments which have not been recognized by Apache Tika (and are therefore not indexed with ePADD), for further review. Note that messages are only flagged in this way if they have also been appraised or processed by ePADD 4.0 or later.

3.6 Exporting Messages to MBOX

Select **Export** from the Appraisal Module menu at the top of the screen to navigate to the Export Screen.

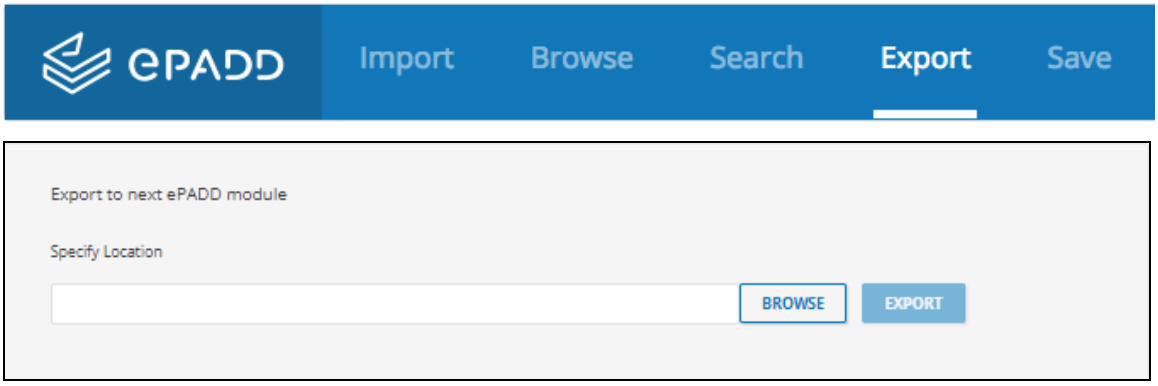


The user can select to export all messages, all non-restricted messages, or just restricted messages. Please note that email exported to MBOX would need to be re-imported into ePADD (and reindexed) in order to benefit from ePADD's functionality in the future.

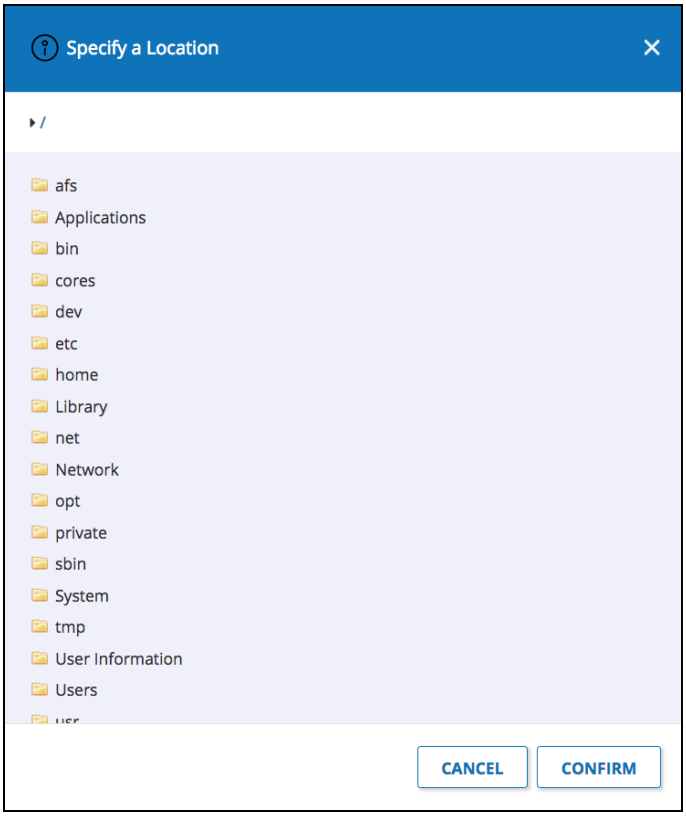
3.7 Exporting Email to the Processing Module

Once you have reviewed the email archive using the tools described above, the final step is to export the appraised email archive (along with whatever supplemental information or annotations you have added) so it can be transferred to an archival repository, and/or processed in another ePADD module.

Select Export from the Appraisal Module menu at the top of the screen to navigate to the Export Screen.



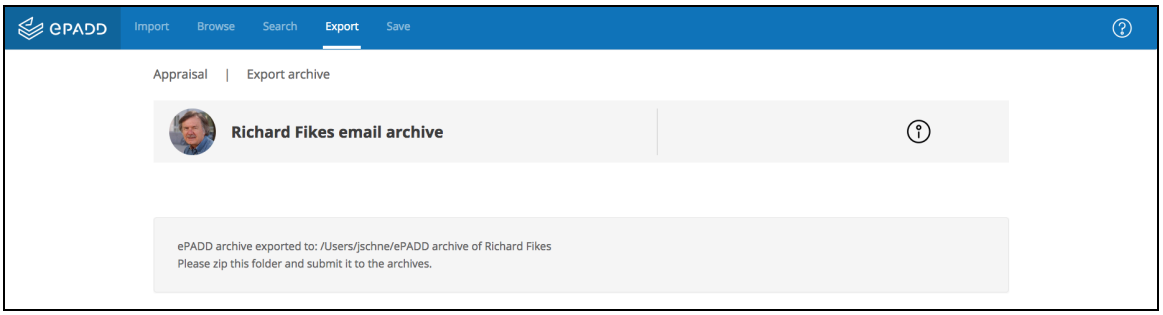
To export the archive to the Processing Module, first select Browse under *Export to Next ePADD Module.*

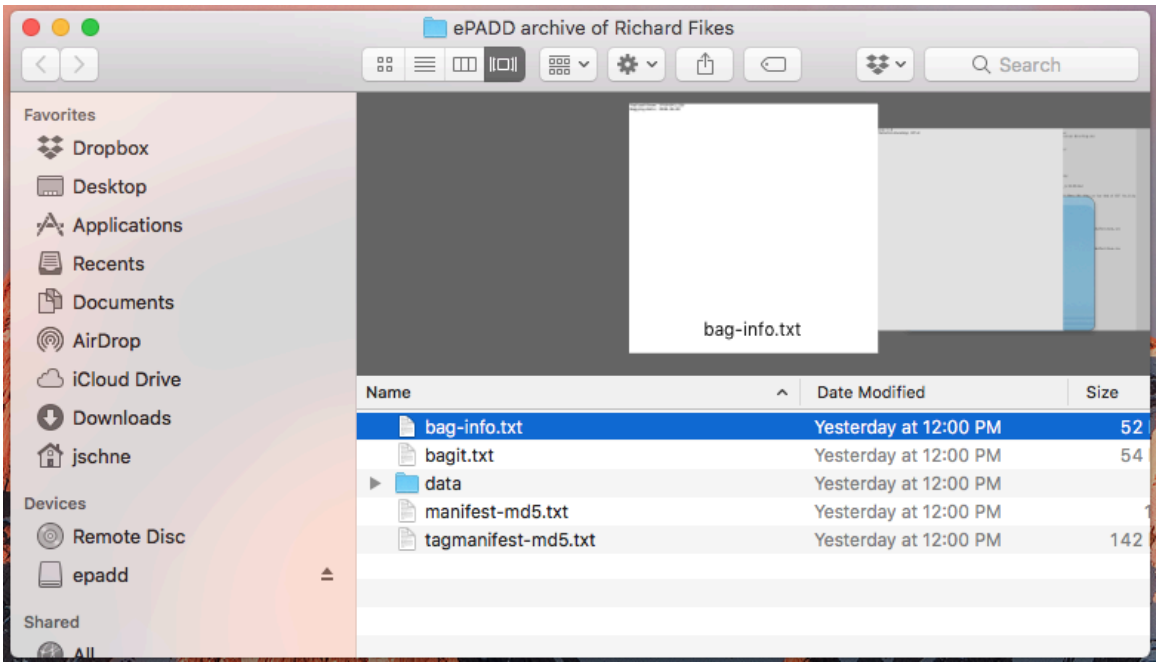


Selecting a folder location to export the appraised email archive

Select the location for ePADD to export the email archive. Please ensure you have enough disk space to save the complete email archive export or the export will fail.

This step may take some time as ePADD assembles the exported email archive and accompanying documentation.





An exported email archive

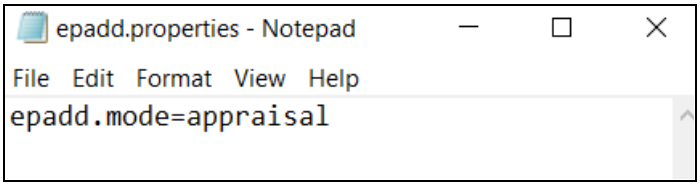
Note to Donor: Once you have confirmed with the archival repository that the transfer is successful, you may wish to delete the ePADD working directories from your workstation. You can do so manually by deleting the following folders under your user directory: 'epadd-appraisal' and 'epadd-settings.'

During the ePADD installation process, a file named *epadd.properties* is created. The file is located in the user's home directory, e.g.

On Windows: c:\users\<username>\
On Mac: /Users/<username>/

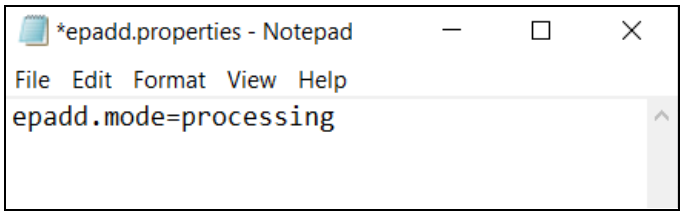
By default, this file contains the following text:

epadd.mode=appraisal



'appraisal' indicates which module or mode ePADD is in; if, for example, you were to restart ePADD it will open in the Apprasial module. To run a different ePADD module, replace 'appraisal' with 'processing', 'discovery' or 'delivery', as in these examples:

- epadd.mode=processing
- epadd.mode=discovery
- epadd.mode=delivery



If it is already running, you must restart ePADD. If no module is specified in the properties files, ePADD will start in its default mode, which is the Appraisal module.

Note that when changing the mode to 'discovery', you also need to start ePADD using the file epadd-discovery-standalone.jar as described elsewhere in this guide (please see section 5.2 Mounting the Discovery Module under a Web Server). Although ePADD will start in discovery mode using epadd.exe or epadd-standalone.jar, it will not completely work as intended.

4 Processing Module

Email archives must initially be ingested to ePADD from the Appraisal module. To work with a collection within the Processing module, it is best practice to first close ePADD, and change the default module using the epadd.properties file (see section 2.2: Installing ePADD in this guide for more details on configuring ePADD). After confirming successful import of files into Processing, you can optionally delete files in Appraisal by deleting the 'epadd-appraisal' folder in your user directory.

If you are planning to run multiple modules on a single workstation, you may also navigate between modules by manually navigating to localhost:9099/epadd/switch-module using the web browser address bar. Again, please note that it is always best practice to close the archive using the 'Close' header menu option, and to fully quit ePADD, before beginning work in another module.

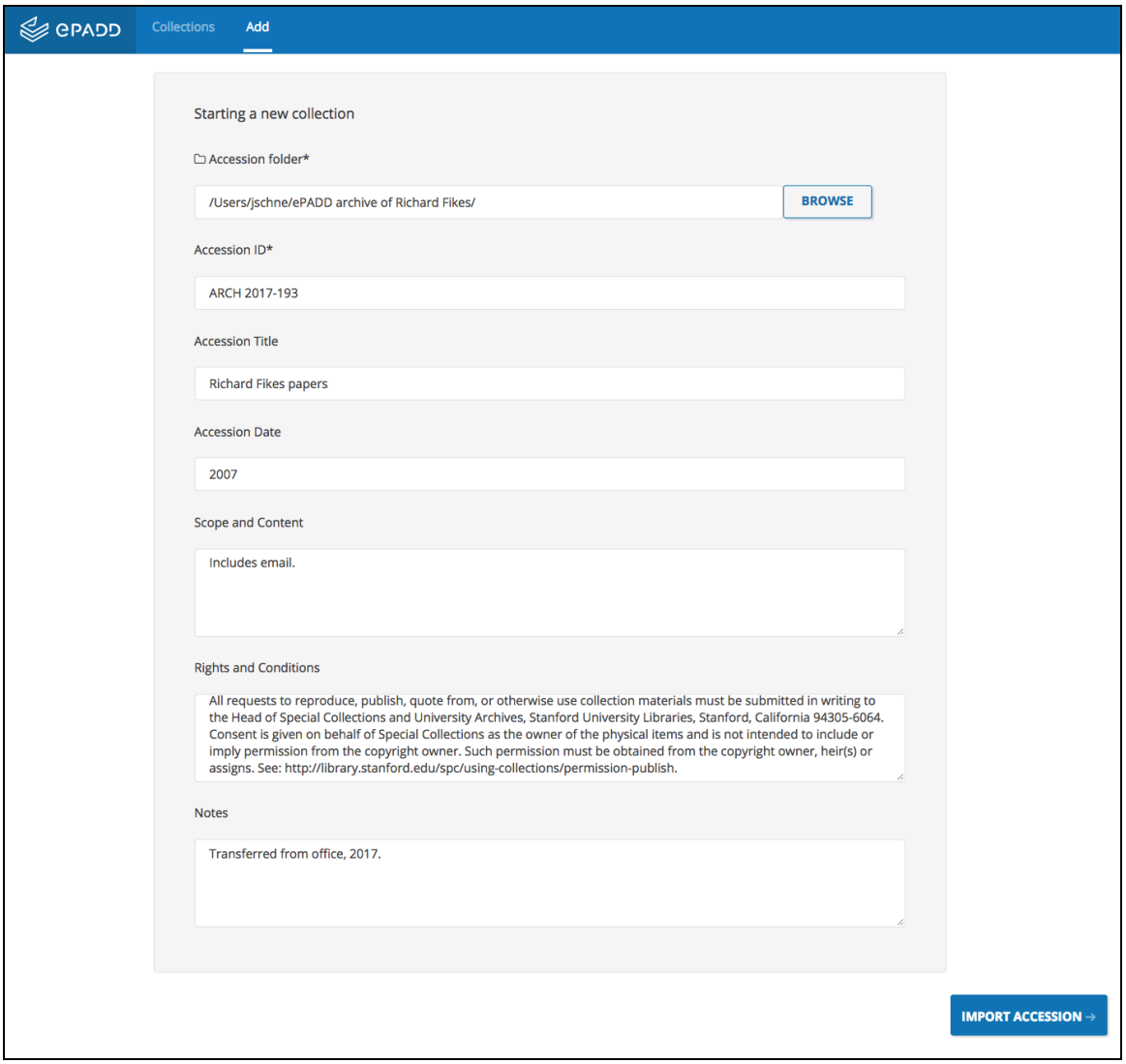
Please note that if you wish to export an unprocessed MBOX file for preservation, this should be your first step upon starting up the Processing Module. Select Export from the navigation bar, then use the Export messages in MBOX format option.

4.1 Importing an Email Archive from the Appraisal Module into the Processing Module and Assigning Accession Metadata

Please note that it is always best practice to fully quit ePADD before beginning work in another module. To work with a collection within the Processing module, it is recommended to first close ePADD, and change the default module using the epadd.properties file (see section 2.2: Installing ePADD of this guide for more details on configuring ePADD).

ePADD uses the concept of accessions and collections. All email exported from Appraisal is considered an accession. To create a new collection within ePADD Processing, select 'Add' from the navigation bar, then use the 'Browse' button to navigate to the folder exported from the Appraisal Module. Assign accession metadata, then import the email archive identified by clicking the 'Import Accession' button.

Users can continue to review and create collection metadata in the Processing Module. Please refer to section 3.3.1 for more information on this process.

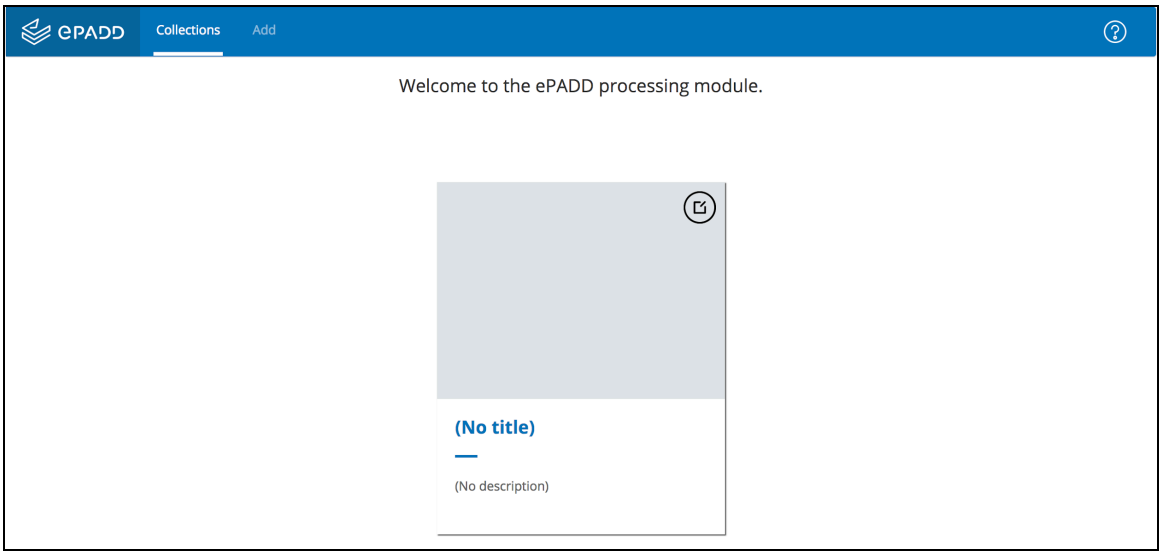


Once you have added an accession to a new collection, you have the ability to add collection metadata.

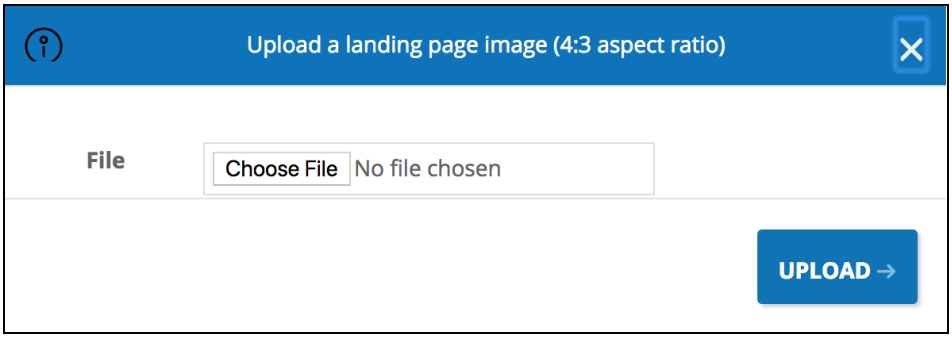
Add an additional accession to an existing collection by selecting the Add Accession button, from the Collection page, below Collection Details. You can also edit accession metadata at any time from the Collection page.

4.2 Assigning Collection Metadata

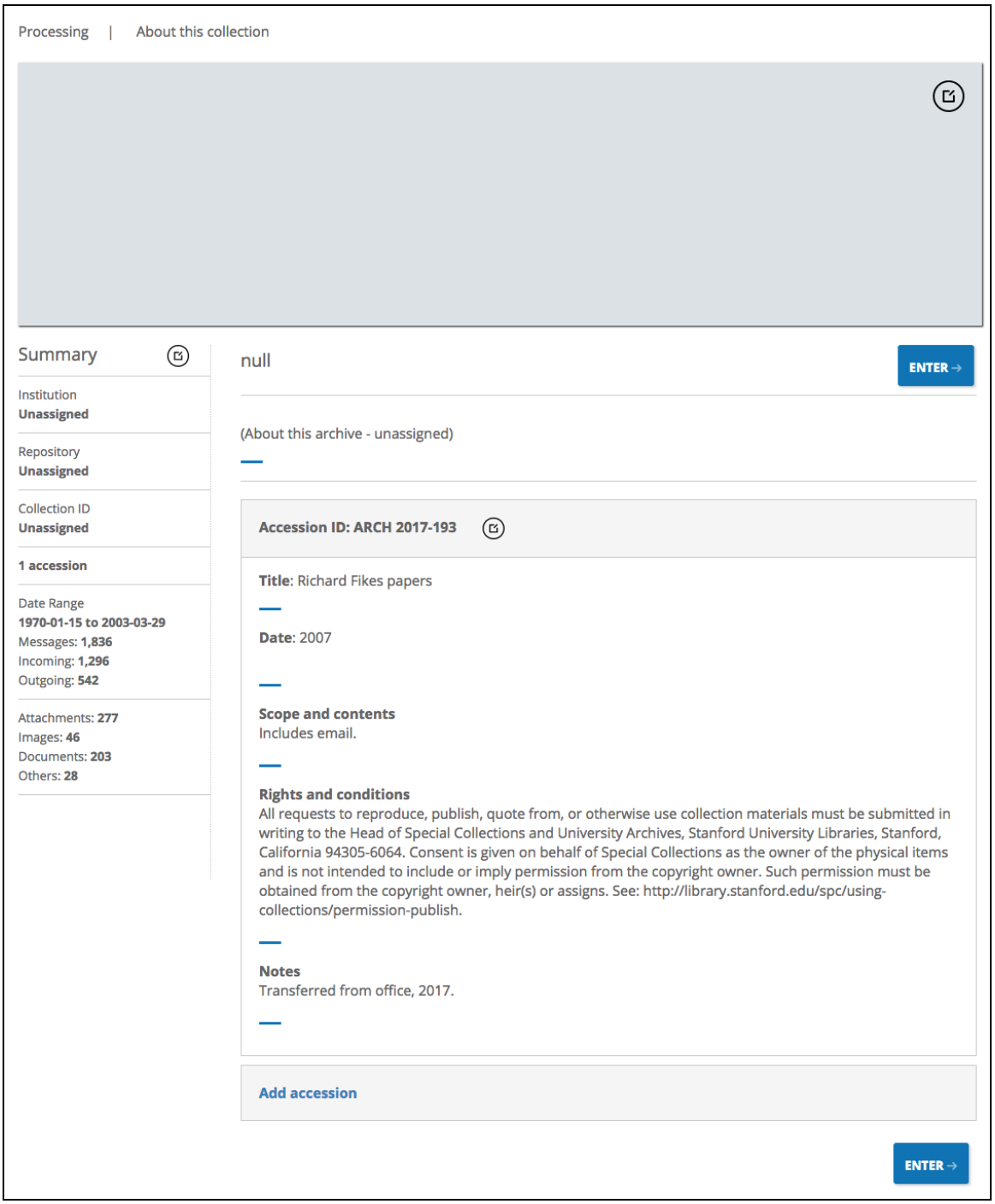
After the accession has been imported, you will be brought to the Browse screen. Close the archive using the close button in the navigation bar to return to the Collections page.



There you will find your accession. Optionally, you can select the edit button to add a landing page image. This image can also be set or changed later.

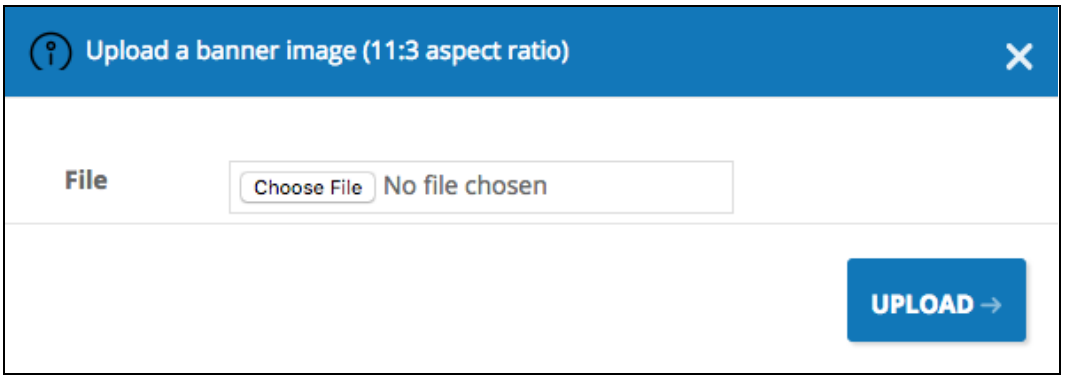


To edit the collection details, click on the image/title.

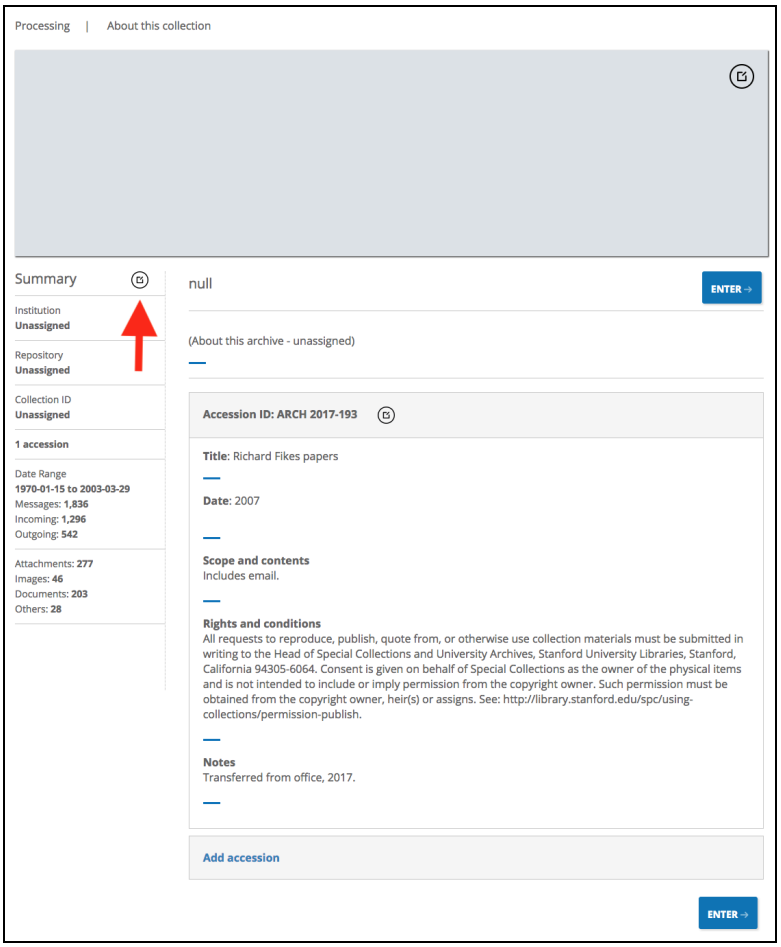


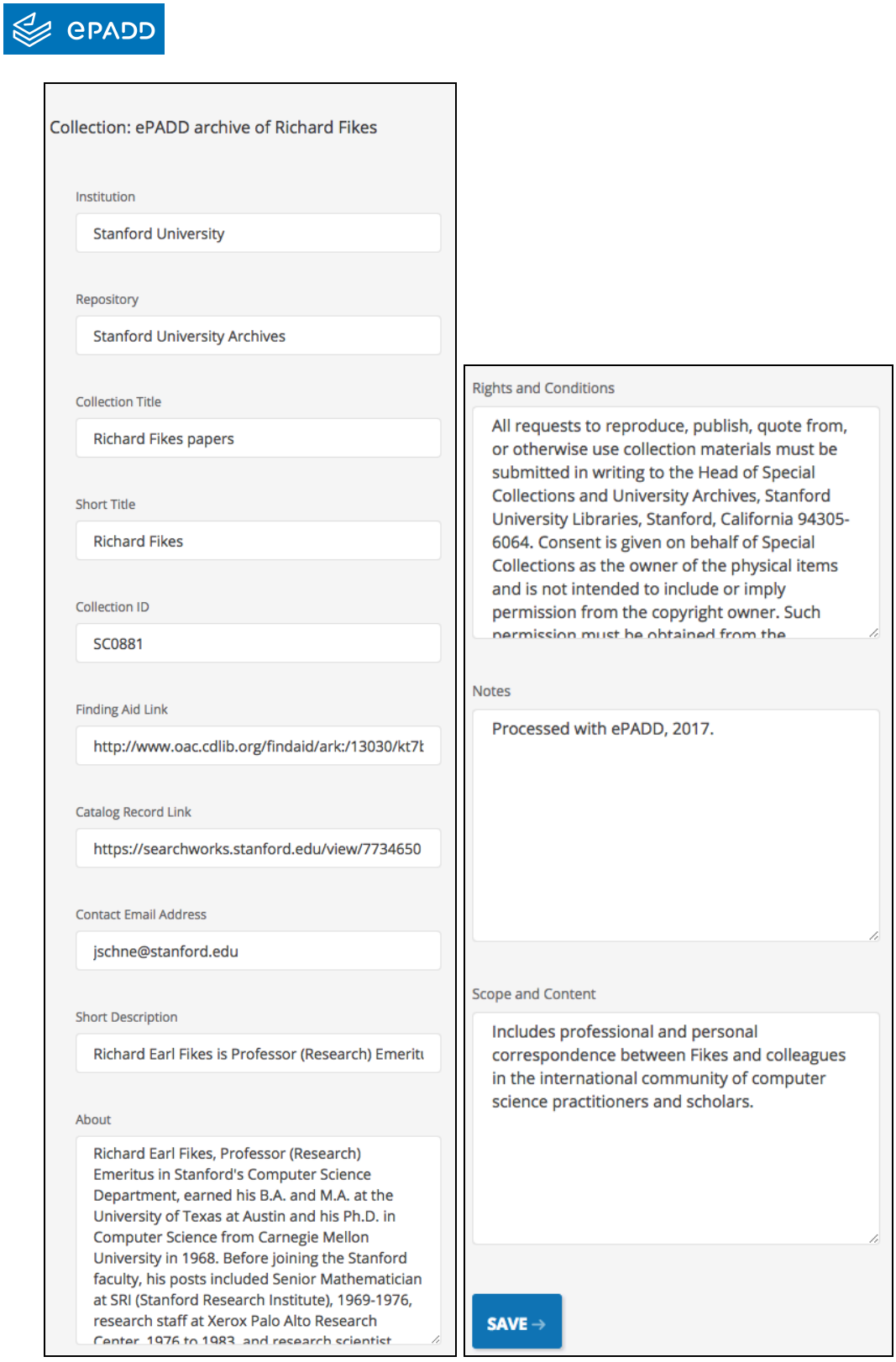
When we are done, this page will match what users will experience in the Discovery and Delivery modules.

Currently, it is lacking a banner image and any collection metadata! Optionally, select the edit button to add a banner image. This image can also be set or changed later.

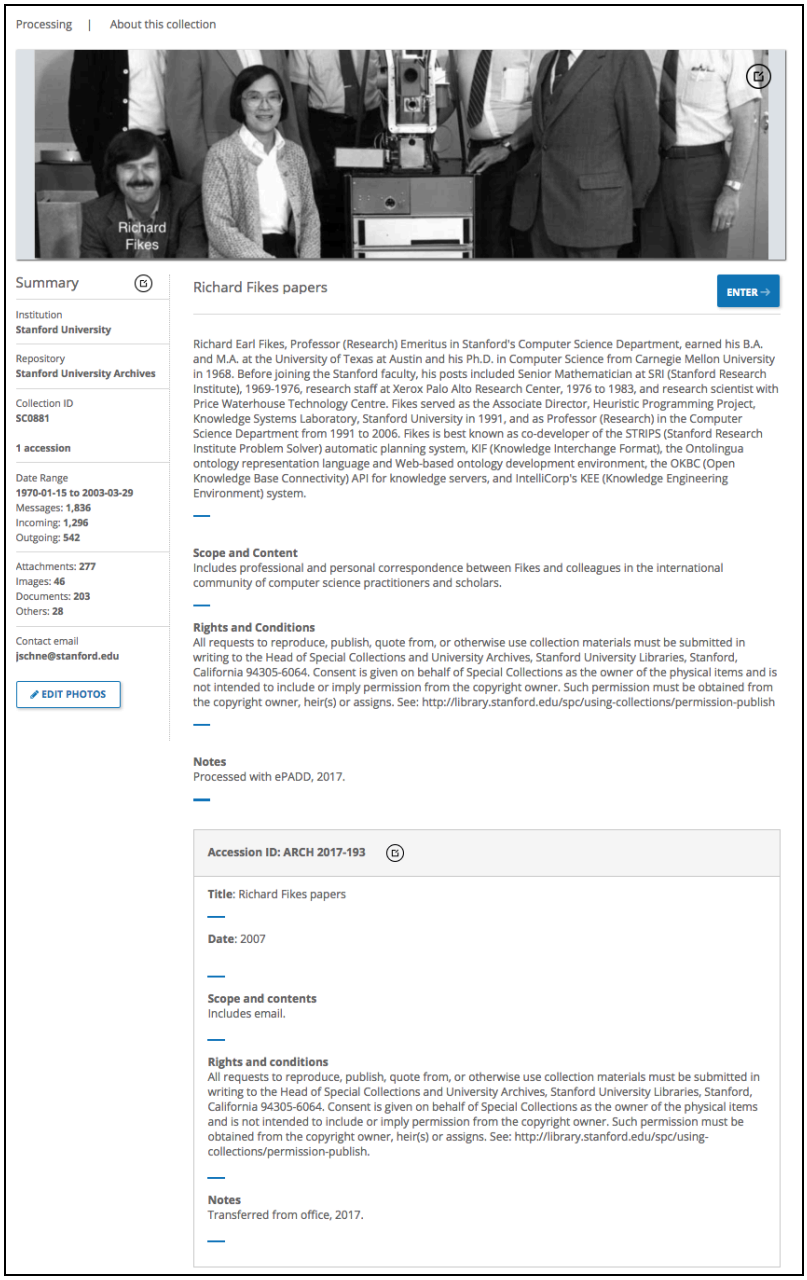


To add collection metadata, select the edit button next to the Summary .





When you have finished adding information, click Save.



Note that if files have been normalized via an Archivematica integration, the number of renamed and normalized files will appear under Collection Details.

4.3 Appraisal Revisited

The **Processing Module** includes all of the review and appraisal functionality of the **Appraisal Module**, as well as additional functionality associated with archival processing. Please refer to **Section 3: Appraisal Module** for information on Reviewing Email (Browsing Email Messages; Editing Correspondents; Performing Entity Analysis, Performing Lexicon

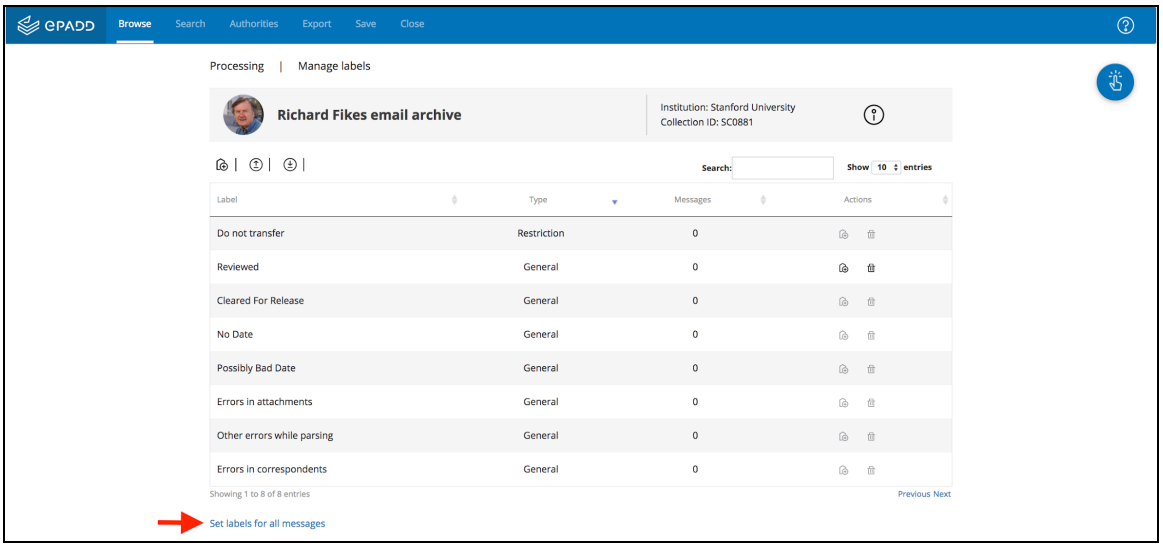
Analysis; and Reviewing Attachments) as well as Appraising Email (Searching Email Messages; Annotating Email Messages; and Restricting Email Messages).

There is no set order to perform the steps above. In fact, different email archives will benefit from different approaches. **However, we do recommend that you begin by verifying the bag checksum. This option can be accessed under the More from the Browse menu**.

The various unique functionalities provided by the Processing Module to assist with archival processing are described below.

4.4 Setting Default Labels

If you wish to reset any labels that were added in Appraisal, or set new default labels, navigate to Labels from the Browse screen and select Set labels for all messages.



The Processing Module contains an additional restriction label not present in the Appraisal Module - 'Transfer to Delivery Only.' When a processed email collection is exported from the Processing Module to the Discovery and Delivery Modules, messages with this label applied will only export to the Delivery Module and not to the Discovery Module.

When processing a large email collection, it is not always possible to carry out appraisal on an entire collection, and a sub-set or smaller collection identified within the collection may be selected as an interim measure. A sub-set of appraised records rather than the collection in its entirety can be transferred to the Delivery module using the ‘permission’ label (We introduced a new type of label: the ‘permission’ label. You have to create a custom label of type ‘permission’ for example ‘Cleared for Transfer to Delivery module’).

4.5 Performing Authority Control

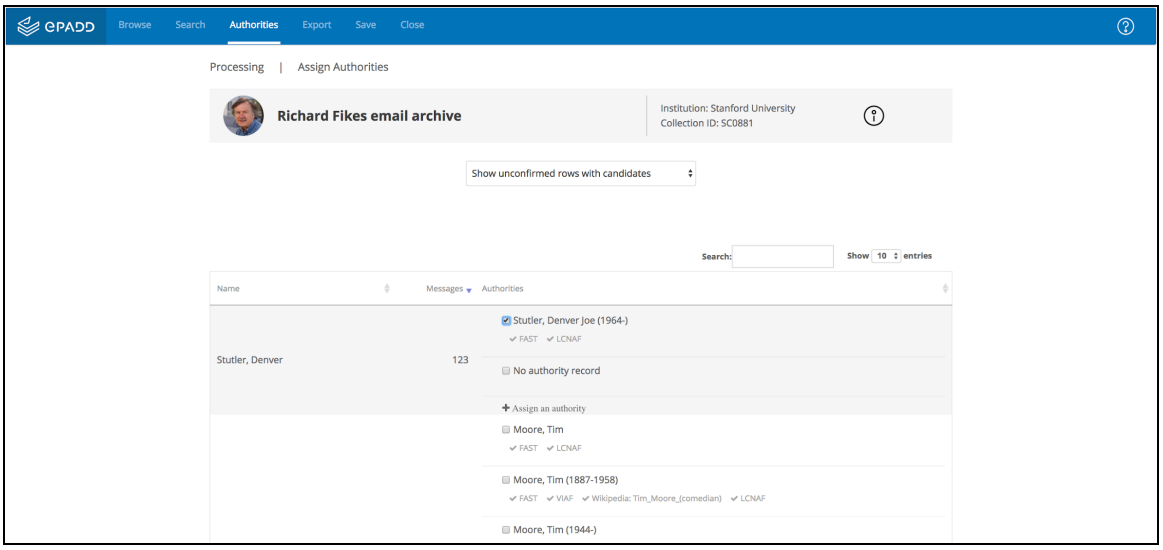
ePADD allows the user to reconcile correspondents and named entities extracted by ePADD with authorized headings (including OCLC FAST headings).

Within the Processing Module, select **Authorities** from the navigation bar to begin this process.

4.5.1 Reconciling Correspondents with Authorized Headings

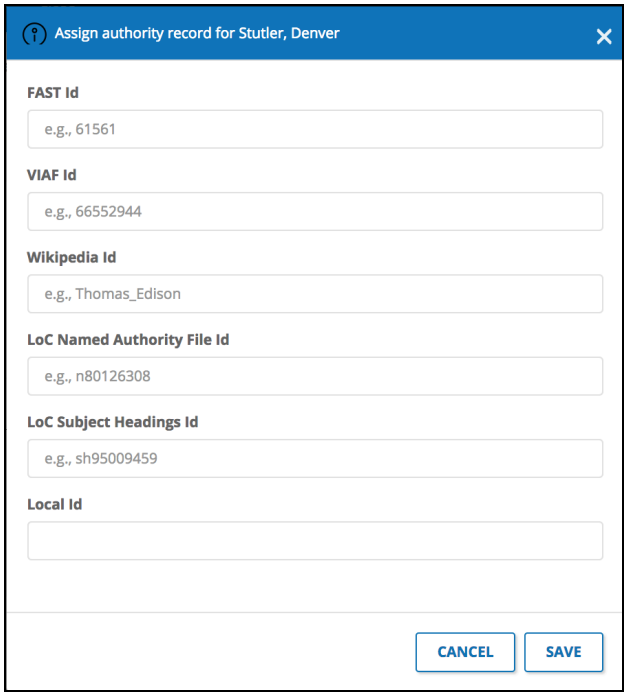
For more information about how ePADD merges correspondents, as well as how to manually override merged correspondents, please see section 3.2.3: Editing Correspondents for further information.

The order of correspondents on the Authorities screen can be relisted alphabetically by person, or by volume of messages exchanged. You can also select how many correspondent entries to view per page using the option box in the top left corner, as well as search for a particular correspondent using the search box at the top of the page.



ePADD searches OCLC FAST for LC Subject Headings, as well as DBPedia, VIAF, and LCNAF. If an image is associated with the related DBPedia record, it can be displayed by hovering over the related DBPedia link.

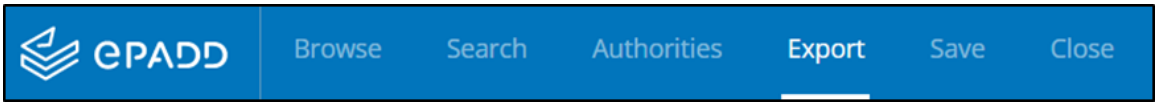
If you are unsatisfied with ePADD's results, you can add an authorized heading ID or local ID using the **+** button.



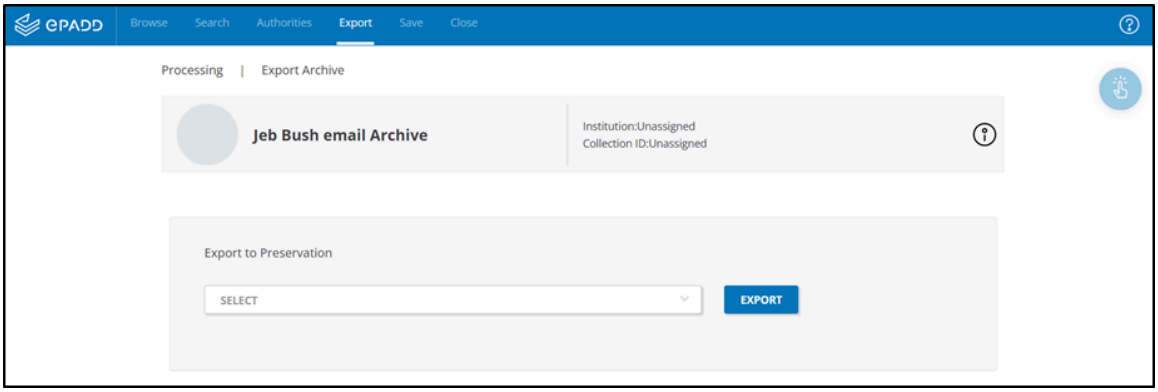
Any actions taken in the Authorities interface are autosaved. Authorized correspondents can be exported as a .CSV file by selecting **Export** from the navigation bar, then using the *Export Correspondents* option. See section 4.11: Exporting Correspondents for more information.

4.6 Exporting Email for Preservation

ePADD can be used to create and export a preservation package or 'preservation ready bag', complete with all email headers, that can be transferred to a digital preservation repository. This can be done, following similar steps, from within ePADD's Appraisal and Processing modules (see section 3.4: Exporting Email for Preservation for exporting email from the Appraisal module). To create the preservation package click 'Export' on the ePADD dashboard:



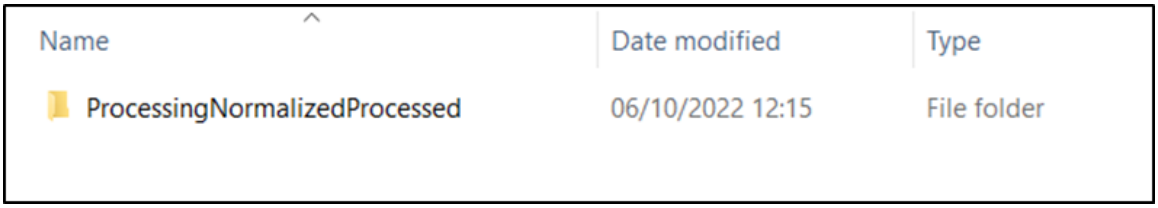
In 'Export to Preservation', select 'exportProcessed' and click 'Export':



Clicking on 'Export' creates a folder – 'ProcessingNormalizedAppraised' – containing the appraised MBOX file(s), which is found via the following file path:

On Windows:
C:\Users\<username>\epadd-appraisal\user\data\exportableAssets\
ProcessingNormalizedProcessed

On Mac: User/<username>/epadd-appraisal/user/data/exportableAssets/
ProcesingNormalizedProcessed

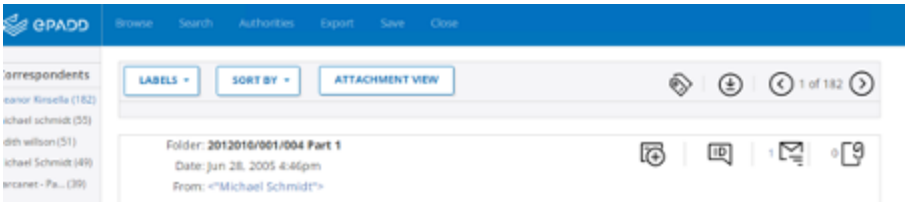


4.7 Downloadable .csv of search results

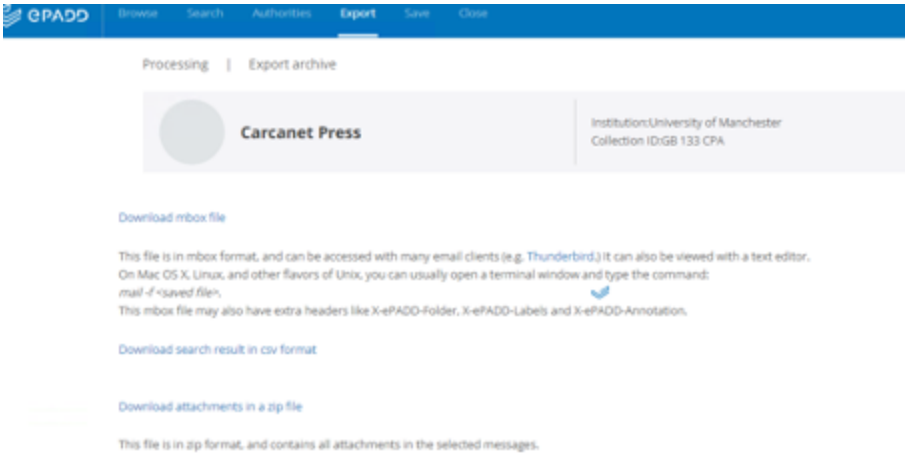
The header metadata for the email selection produced by the results of any search, generated by any means, (search, advanced search, from the correspondents list, from the lexicon) can be downloaded as a CSV file.

Instructions

Select the 'download messages as MBOX' option on the selected search results page:

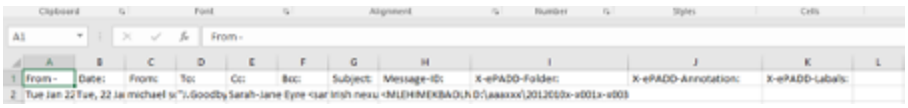


The following options are presented on the next screen, where you can select 'download search result in CSV format':



The CSV will then download, consisting of the following header data fields:

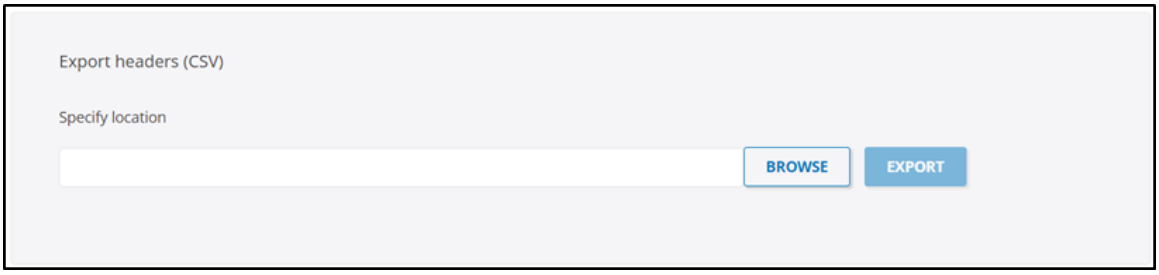
- Date
- From
- To
- CC
- BCC
- Subject
- Message ID
- Folder
- Annotation
- Label



4.8 Exporting Message Headers

Message headers can be exported from Processing as a .CSV file to aid in social network analysis using the *Export Headers (CSV)* option. Exported fields include Date, From (resolved name), from (email address), To (resolved name), To (email address), CC (resolved name), CC (email address), BCC (resolved name), BCC (email address).

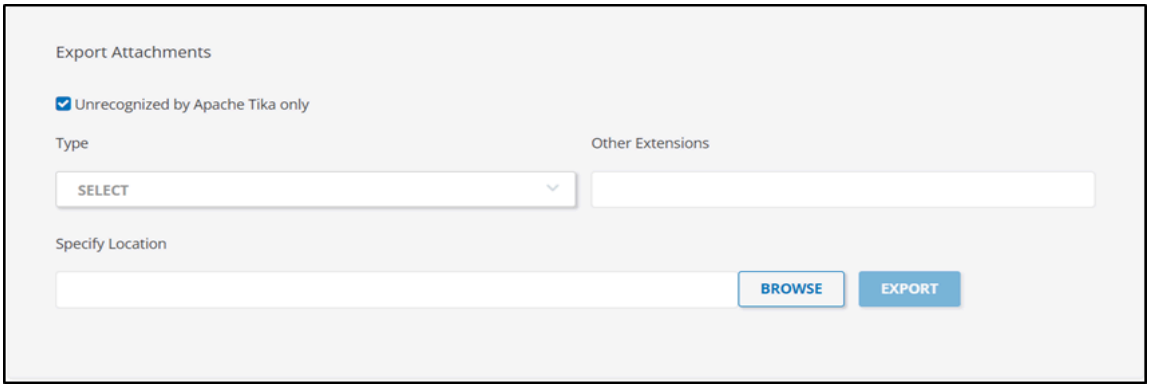
Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen, where you will find the section entitled 'Export Headers (CSV):



4.9 Exporting Message Attachments

Users can also choose to export attachments according to various search criteria. The user can indicate whether they would like to export only those attachments which have not been recognized by Apache Tika (and are therefore not indexed with ePADD) for further review. Note that messages are only flagged in this way if they have also been appraised or processed by ePADD 4.0 or later.

Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen, where you will find a section entitled 'Export Attachments':



4.10 Exporting to MBOX

Users can export messages to MBOX -- the user can select to export all messages, all non-restricted messages, or just restricted messages. Please note that email exporting to MBOX will need to be re-imported into ePADD (and reindexed) in order to benefit from ePADD's functionality in the future.

Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen, where you will find a section entitled 'Export Messages in MBOX Format':

4.11 Exporting Entities

Users can export recognized entities. The user can select whether they would like to export all entities or just particular entity types. Entities are exported as CSV and include the following fields: Entity name, number of related messages, start and end dates for related messages, entity type, and score (whether there is a direct match with an entity in DBPedia).

Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen, where you will find a section entitled 'Export Entities':

4.12 Exporting Correspondents

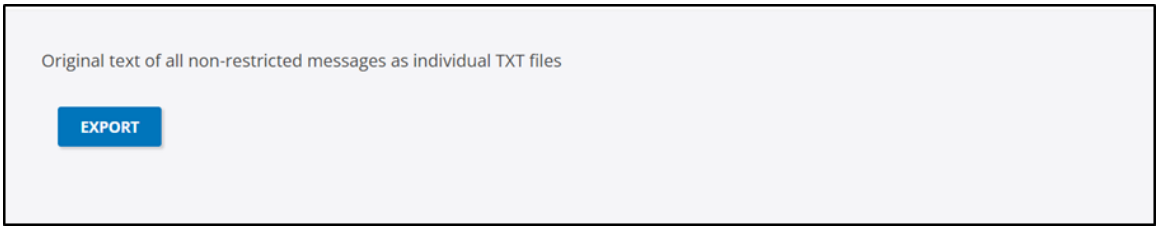
Users can export correspondents from the email archive. Users can choose to export only correspondents that have been confirmed via the Authorities interface (which include related unique identifiers), or only unconfirmed correspondents. Unconfirmed correspondents are exported as CSV and include the following fields: Correspondent name, number of sent and received messages, number of mentions, and start and end dates for related messages.

Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen, where you will find a section entitled 'Export correspondents':

4.13 Exporting Original Text of Non-restricted Messages as Individual Text Files

Users can export all message text (subject line and message body) as individual TXT files to support topic modeling and other methods of text analysis.

Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen, where you will find the relevant section:

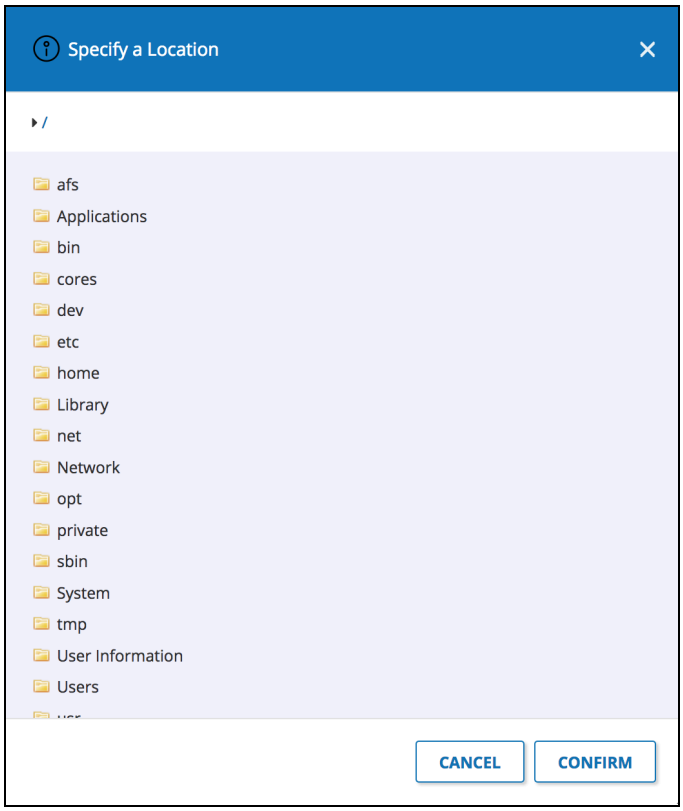


4.14 Exporting Email to the Delivery and Discovery Modules

Once you have reviewed the email archive using the tools described above, the final step is to export the processed email archive (along with whatever supplemental information or annotations you have added) for display in the Discovery and/or Delivery Modules.

Select **Export** from the Processing Module menu at the top of the screen to navigate to the Export Screen.

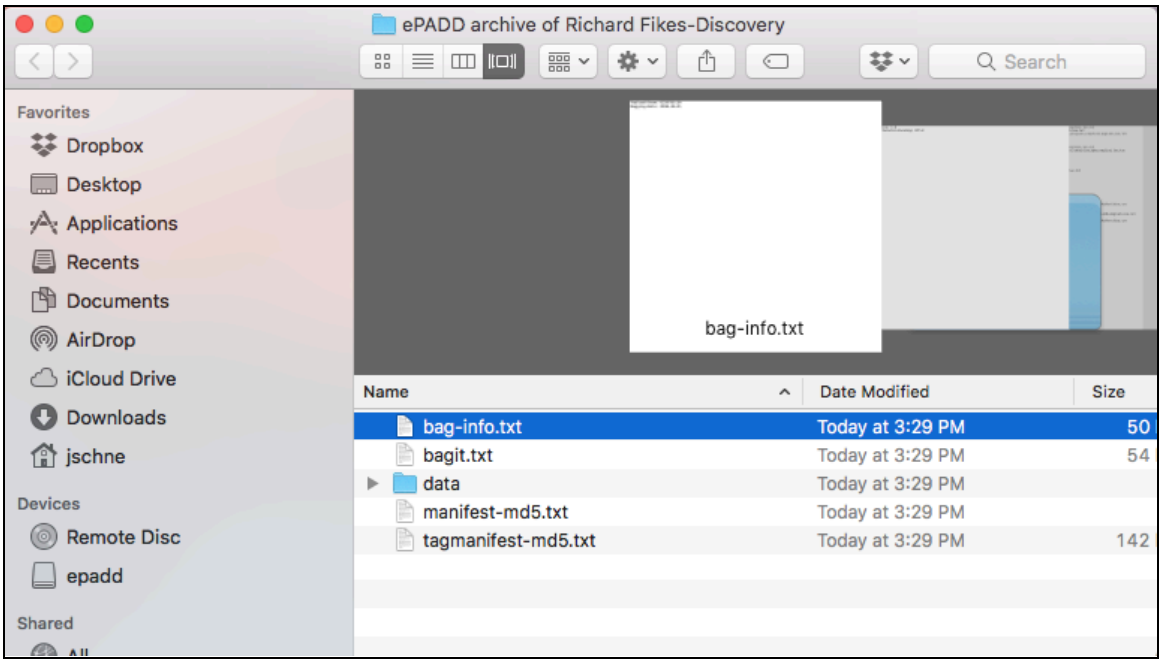
To export the archive to the Discovery and Delivery modules, first use the Browse button to select the location you wish to export under *Export to next ePADD module,* then click the confirm button.

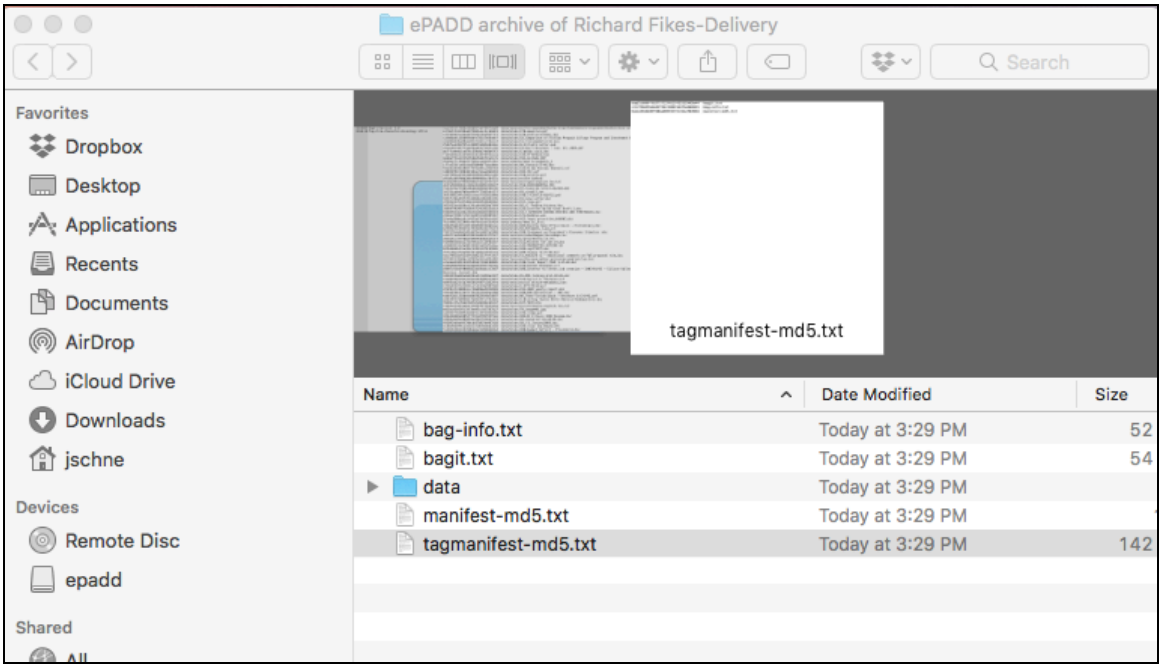


Selecting a folder in which to export the email archive

Please ensure you have enough disk space to save the complete email archive export or the export will fail.

This step may take some time as ePADD assembles the exported email archive and accompanying documentation.





Exporting the processed email archive from the Processing Module creates two folders: one to support Discovery, and one to support Delivery

5 Discovery Module

The Discovery Module is designed to run under a web-server, and allow remote users to search redacted header info and extracted entities while limiting full-text access to the materials.

5.1 Testing the Discovery Module

To test the Discovery module on your local workstation, copy the folder 'Email archive of <name>-Discovery,' exported through the Processing Module to the 'epadd-discovery' folder in your user directory. (Please see section 2: Getting Started for more information about locating your user directory). If an 'epadd-discovery' folder does not already exist within your user directory, create that folder.

Now run ePADD and navigate to the Discovery module within the settings folder. At this point, you may wish to reset the 'reviewed' status for all messages that may have been marked as 'reviewed' in the Processing Module. To do this, navigate to the Settings Menu, and select the 'Set Default Actions' button.

5.2 Mounting the Discovery Module under a Web Server

Please follow the steps below to mount the ePADD Discovery Module under a web server. As the software is still in development, please contact us **(epadd_project@stanford.edu)** to troubleshoot any issues and/or confirm your successful installation.

1) Install the **Java runtime environmen**t (JRE 11+) on the machine that is going to mount the web server.

2) Create a file called epadd.properties. It must be in plain text and stored in the user's home directory (typically c:\users\<username>\).

This file should contain the following text:

epadd.mode=discovery

3) Copy the exported Discovery version of the archive folders (which should be titled *ePADD archive of <name>-Discovery)* into ~/epadd-discovery of the home directory for the account that will be running the ePADD discovery module / web server.

Upon copying the files you should see the following directories, including subdirectories and files:

$ ls ~/epadd-discovery

ePADD archive of <name>-Discovery

Important: Please ensure that you do not copy the Delivery version of the archive. It MUST be the Discovery version, as the Delivery version has unredacted information!

4) Download and install the Discovery Module distribution file (**epadd-discovery-standalone.jar**). The web server is bundled with this distribution.

5) Start ePADD with:

java -Djava.io.tmpdir=/tmp -Depadd.mode.discovery=1 -Dnobrowseropen= -Xmx4g -jar epadd-discovery-standalone.jar
Please adjust the -Xmx option to assign as much memory you want, depending on the size of the archive(s).

Please also note that the java.io.tmpdir should be set to a directory where temporary files can be stored by ePADD. (java.io.tmpdir=/tmp will not work on Windows systems).

Please also ensure that the 9099 port is accessible from outside and will remain open.

Caution: Please ensure that this directory is not periodically cleared by utilities like tmpwatch.

6) You can now open ePADD application in your internet browser by typing the URL
http://{machine-name}:9099/

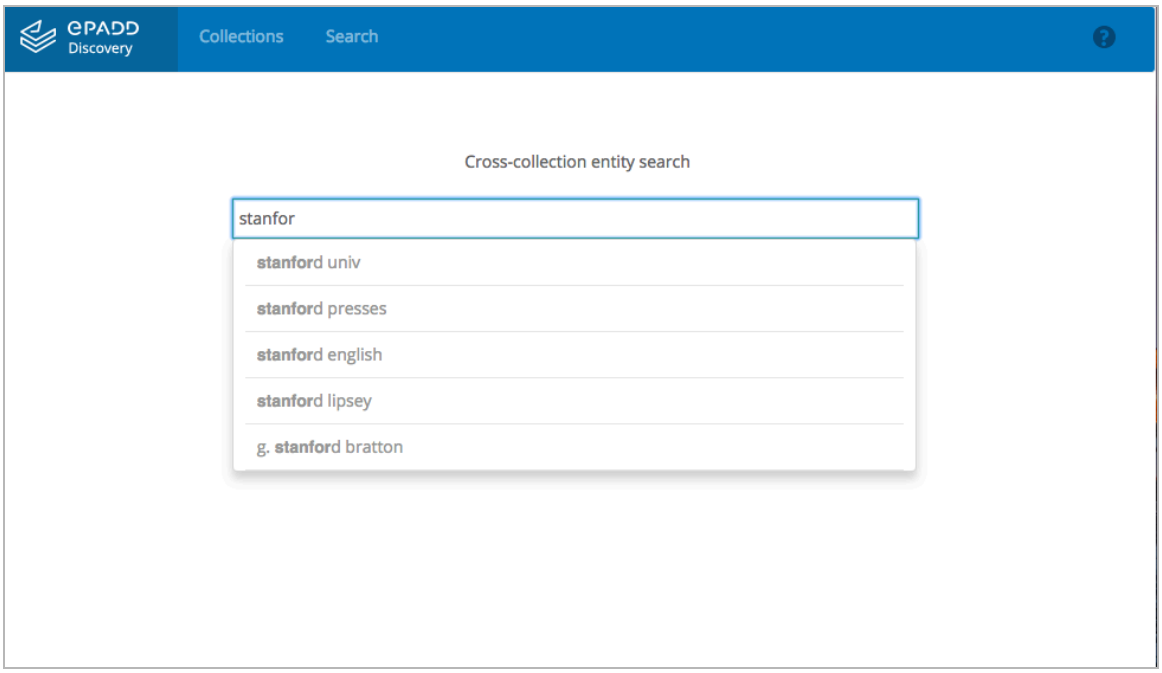
Where {machine-name} is the publicly visible name/IP of the machine where the epadd-discovery module is running.

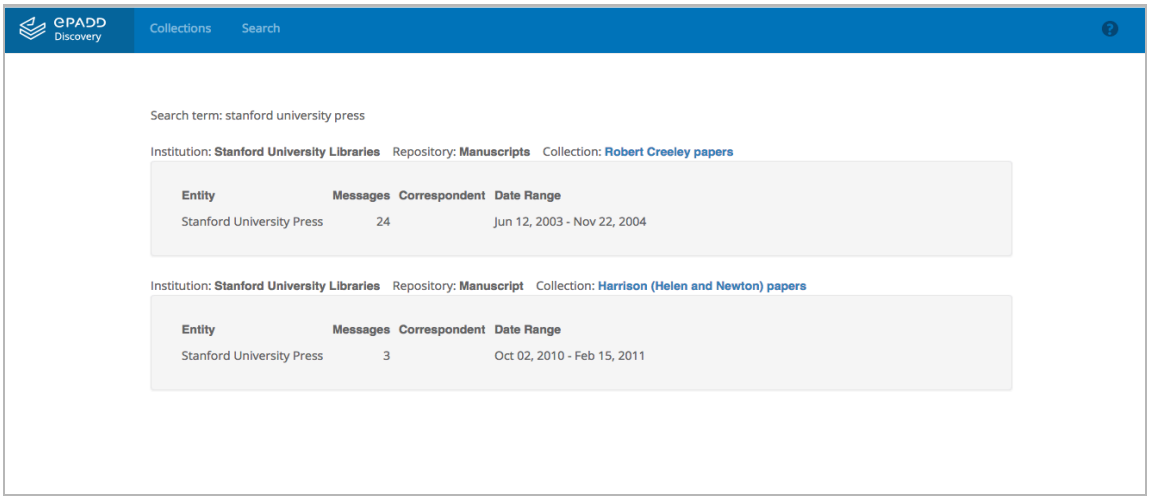
7) You may wish to reset the 'reviewed' status for all messages that may have been marked as 'reviewed' in the Processing Module. To do this, navigate to the Settings Menu, and select the 'Set Default Actions' button.

Note: Please contact the ePADD development team (epadd_project@stanford.edu) for help troubleshooting installation of the Discovery Module.

5.3 Searching Entities Across Collections

ePADD supports cross-collection entity search. Navigate to Search from the Discovery navigation bar to access cross-collection entity search.



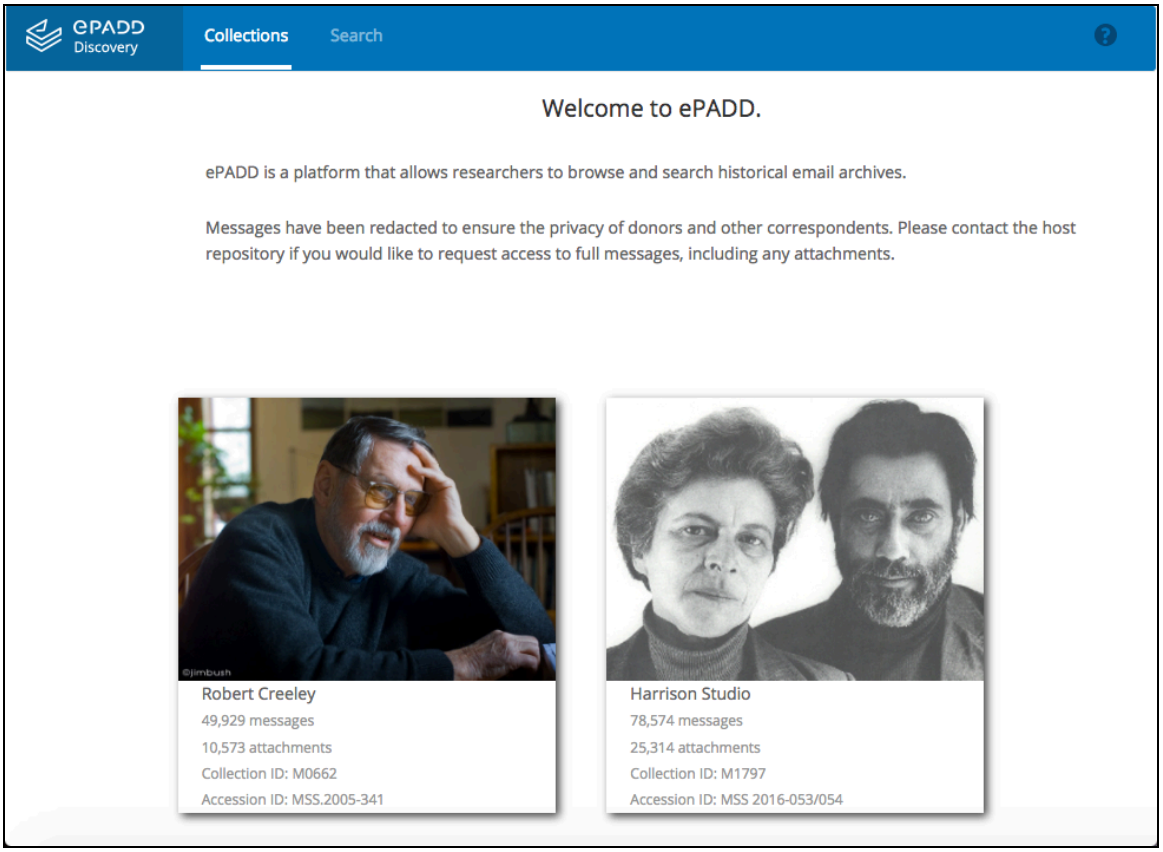


Cross-collection entity search supports autocomplete. The user can navigate from search results to the related collection.

5.4 Navigating to a Collection

The Collections menu provides a user with access to the redacted email of any processed collections.

The user selects the collection they wish to access by clicking on the related image.



Selecting a Collection

The next screen provides basic details of the collection. A user can proceed with viewing the collection by selecting the 'Enter' button.

ePADD

Robert Creeley papers

Collection Details

Institution
Stanford University

Repository
Special Collections & University Archives

Collection ID
M0662

Accession ID
ACCN 2005-182

Date Range
1995-03-05 to 2005-03-23
8 messages undated

Messages: 1,680
Incoming: 868
Outgoing: 15

Attachments: 245
Images: 34
Documents: 84
Others: 127

Contact email
specialcollections@stanford.edu

About Robert Creeley

The Robert Creeley Papers document the life work of a leading American poet of the 20th century, one of the core members of the "Black Mountain School." They also document several important movements in American poetics in the second half of the century. The papers include Creeley's personal and professional correspondence, journals, business records, personal mementos, clippings, artwork, and other documents generated and collected by him from 1950 to 1997.

Rights and Conditions
Property rights reside with the repository. Literary rights reside with the creators of the documents or their heirs. To obtain permission to publish or reproduce, please contact the Public Services Librarian of the Dept. of Special Collections.

Notes
Additional Creeley materials can be found in M0553 and M0347.

ENTER →

Finding Aid

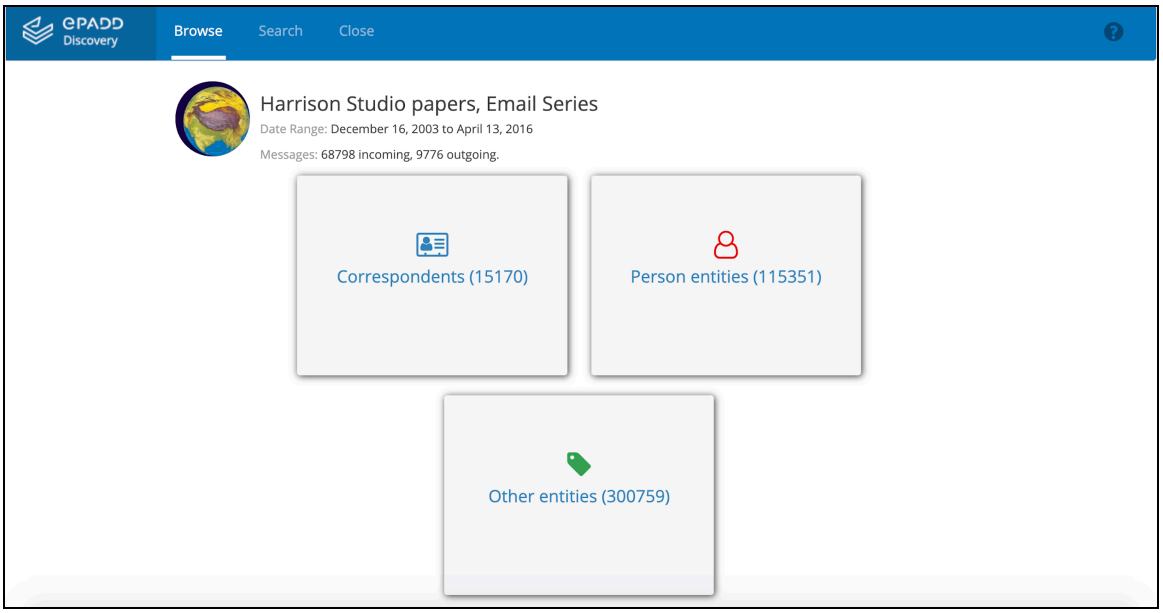
Catalog Record

Email messages in the ePADD Discovery Module have been redacted to ensure the privacy of donors and other correspondents. Please contact the host repository if you would like to request access to full messages, including any attachments.

Access a Collection by clicking the 'ENTER' button.

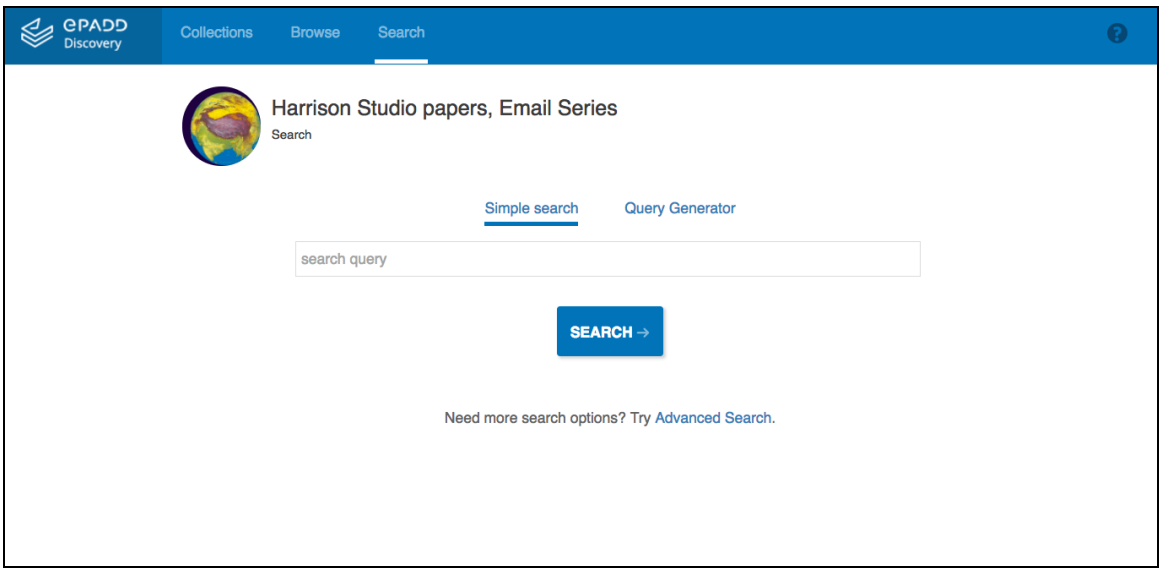
5.5 Browsing Email

The Browse Menu provides a user with access to several options to browse a redacted email archive, including by Correspondent as well as by extracted entity type (persons, organizations, and locations).



Browse Menu in the Discovery Module

5.6 Searching Email



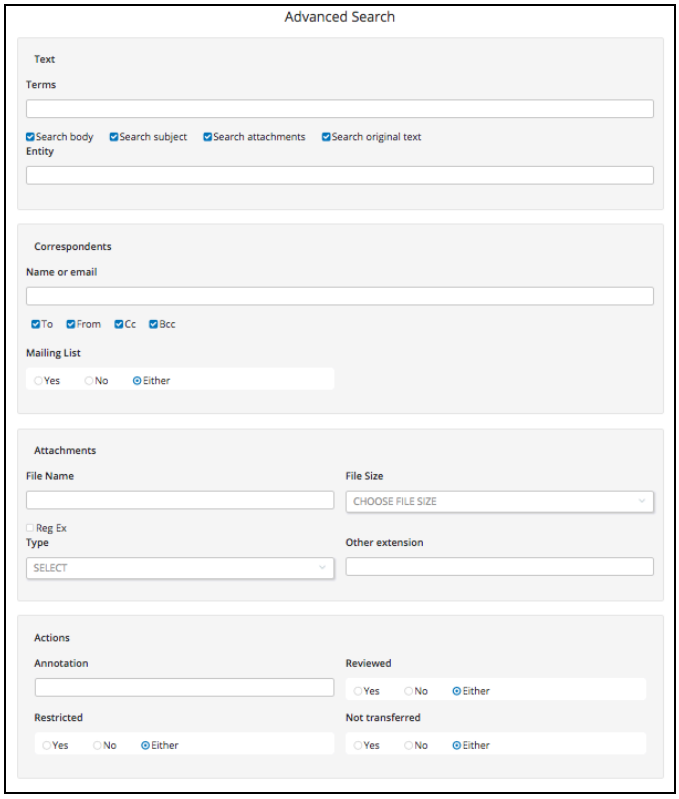
To search the email archive, navigate to the Search screen.

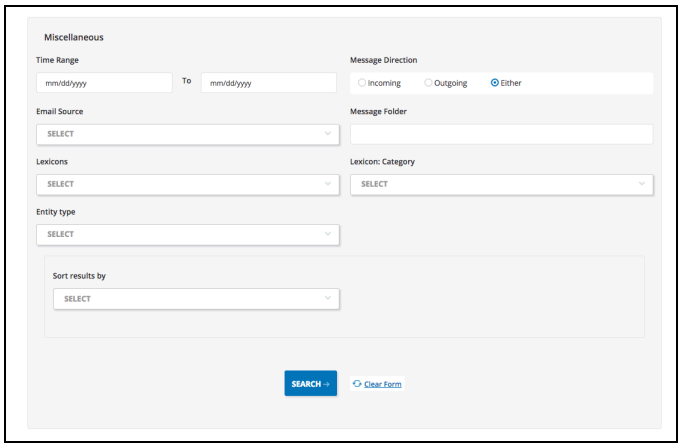
Beginning with v2.0, ePADD supports three types of search: simple search, advanced search, and the query generator.

To use the simple search feature, simply type in your query into the simple search box available on the initial search screen. This will search across message headers (correspondents and subjects) and message text, as well as the text of all document attachments.

Once you have entered your term(s) and made your selection, click the *Search* button. The search string will appear highlighted on the message browse screen.

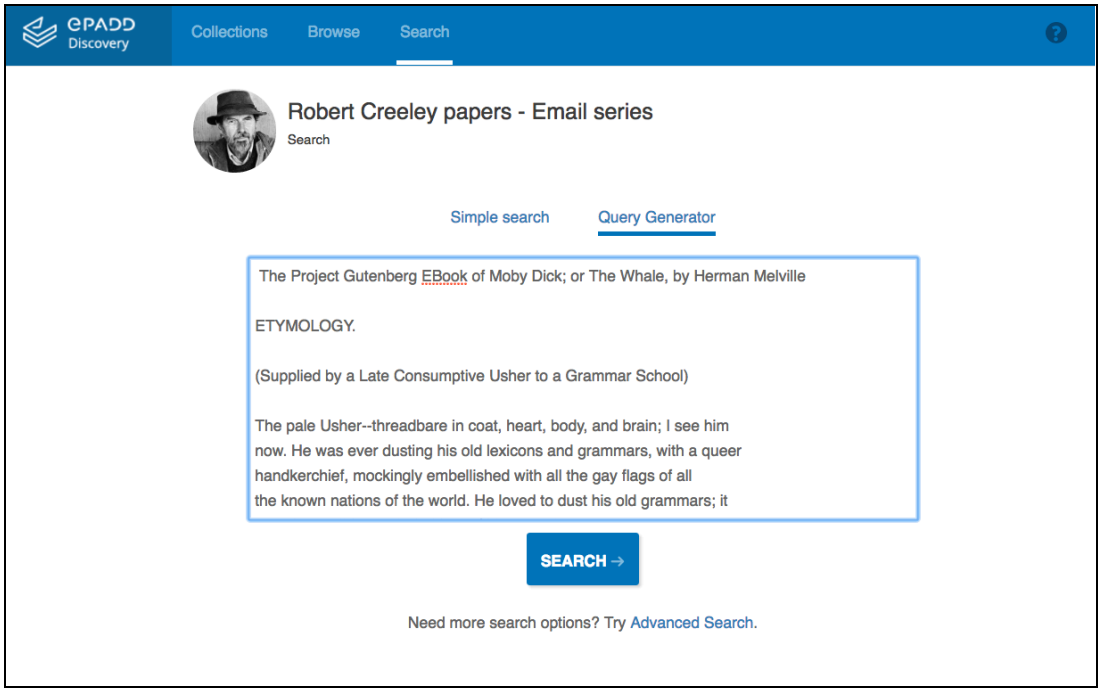
Advanced Search enables the user to to generate complex searches of email text, correspondents, attachments, and any actions previously taken on the messages, as well as a variety of other parameters.



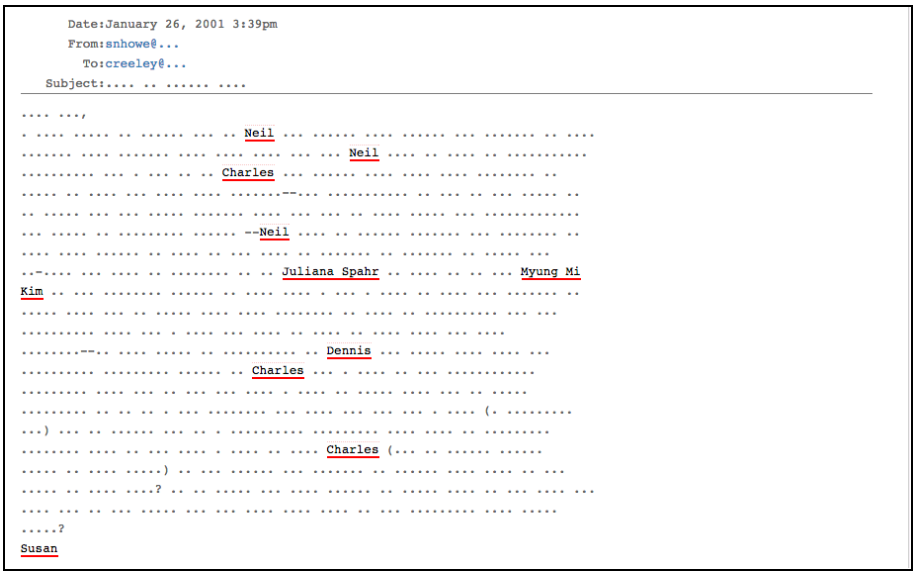


Search Tips can be accessed by selecting the button in the top right corner of the *Advanced Search* screen.

ePADD also allows a user to search the email archive for entities contained in a given block of provided text, using the *Query Generator*. Any matching terms will be highlighted on the results page.



Search options within the Discovery Module



Discovery Module search results, displaying only redacted headers and extracted entities

Query generator in action

6 Delivery Module

Delivery Module Navigation Bar

The Delivery Module enables archival repositories to provide moderated full-text access to unrestricted email archives within a reading room.

6.1 Importing the Email Archive into the Delivery Module

In order for ePADD to recognize the email archive, you must first ensure it is in the proper location within your user directory. (See section 2: Getting Started for more information about locating your user directory). If an 'epadd-delivery' folder does not already exist within your user directory, create that folder, and add the appropriate email archive (which should be a folder titled 'Email Archive of <Archive Owner>-Delivery') to that location.

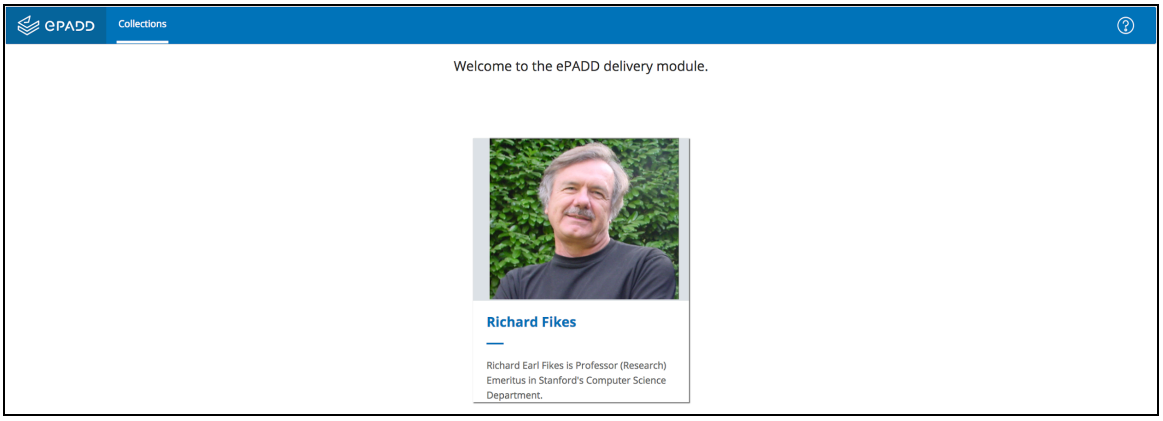
Now run the ePADD software, and select the Delivery Module from the Settings Menu. At this point, you may wish to reset the 'reviewed' status for all messages that may have been marked as 'reviewed' in the Processing Module.. To do this, navigate to the Settings Menu under Browse and select the 'Set Default Actions' button.

The Delivery module was modified to display a redacted version of email addresses in the header metadata, similar to the display available in the Discovery module, to accommodate the data protection requirements on giving access to personal data.

6.2 Navigating to a Collection

The Collections menu provides a user with moderated access to the email archive of any processed collections.

The user selects the collection they wish to access by clicking on the related image.



Selecting a Collection

The next screen provides basic details of the collection. A user can proceed with viewing the collection by selecting the 'Enter' button.

ePADD

Delivery | About this collection

Summary

Institution
Stanford University

Repository
Stanford University Archives

Collection ID
SC0881

1 accession

Date Range
1970-01-15 to 2003-03-29
Messages: **1,836**
Incoming: **1,296**
Outgoing: **542**

Attachments: **277**
Images: **46**
Documents: **203**
Others: **28**

Contact email
jschne@stanford.edu

Richard Fikes papers

ENTER →

Richard Earl Fikes, Professor (Research) Emeritus in Stanford's Computer Science Department, earned his B.A. and M.A. at the University of Texas at Austin and his Ph.D. in Computer Science from Carnegie Mellon University in 1968. Before joining the Stanford faculty, his posts included Senior Mathematician at SRI (Stanford Research Institute), 1969-1976, research staff at Xerox Palo Alto Research Center, 1976 to 1983, and research scientist with Price Waterhouse Technology Centre. Fikes served as the Associate Director, Heuristic Programming Project, Knowledge Systems Laboratory, Stanford University in 1991, and as Professor (Research) in the Computer Science Department from 1991 to 2006. Fikes is best known as co-developer of the STRIPS (Stanford Research Institute Problem Solver) automatic planning system, KIF (Knowledge Interchange Format), the Ontolingua ontology representation language and Web-based ontology development environment, the OKBC (Open Knowledge Base Connectivity) API for knowledge servers, and IntelliCorp's KEE (Knowledge Engineering Environment) system.

Scope and Content
Includes professional and personal correspondence between Fikes and colleagues in the international community of computer science practitioners and scholars.

Rights and Conditions
All requests to reproduce, publish, quote from, or otherwise use collection materials must be submitted in writing to the Head of Special Collections and University Archives, Stanford University Libraries, Stanford, California 94305-6064. Consent is given on behalf of Special Collections as the owner of the physical items and is not intended to include or imply permission from the copyright owner. Such permission must be obtained from the copyright owner, heir(s) or assigns. See: http://library.stanford.edu/spc/using-collections/permission-publish

Notes
Processed with ePADD, 2017.

Accession ID: ARCH 2017-193

Title: Richard Fikes papers

Date: 2007

Scope and contents
Includes email.

Rights and conditions
All requests to reproduce, publish, quote from, or otherwise use collection materials must be submitted in writing to the Head of Special Collections and University Archives, Stanford University Libraries, Stanford, California 94305-6064. Consent is given on behalf of Special Collections as the owner of the physical items and is not intended to include or imply permission from the copyright owner. Such permission must be obtained from the copyright owner, heir(s) or assigns. See: http://library.stanford.edu/spc/using-collections/permission-publish.

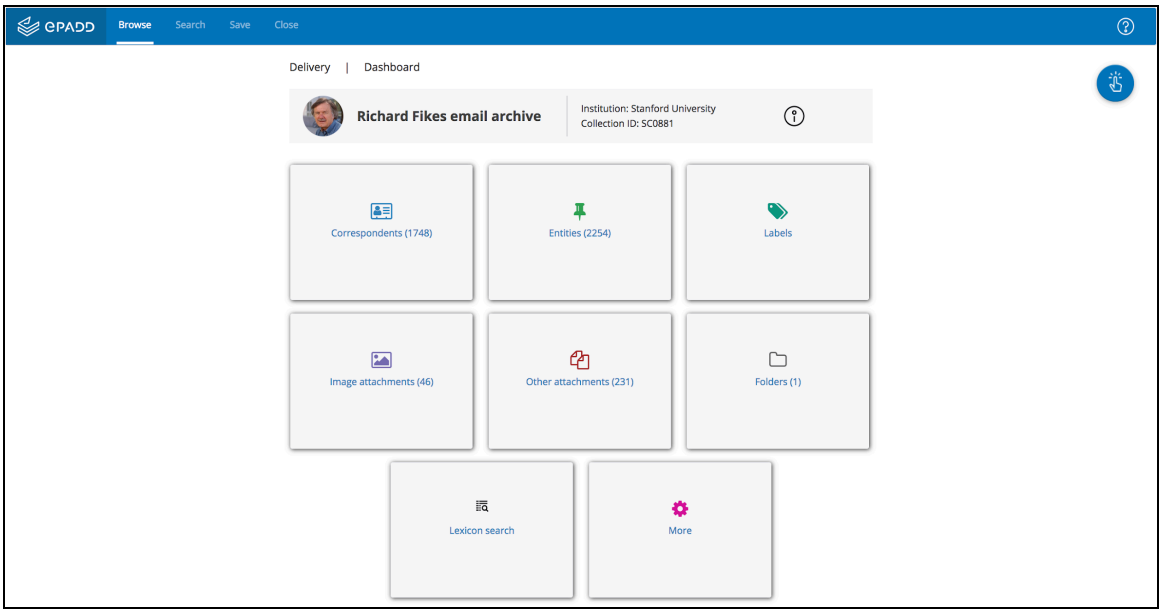
Notes
Transferred from office, 2017.

Finding Aid Catalog Record

ENTER →

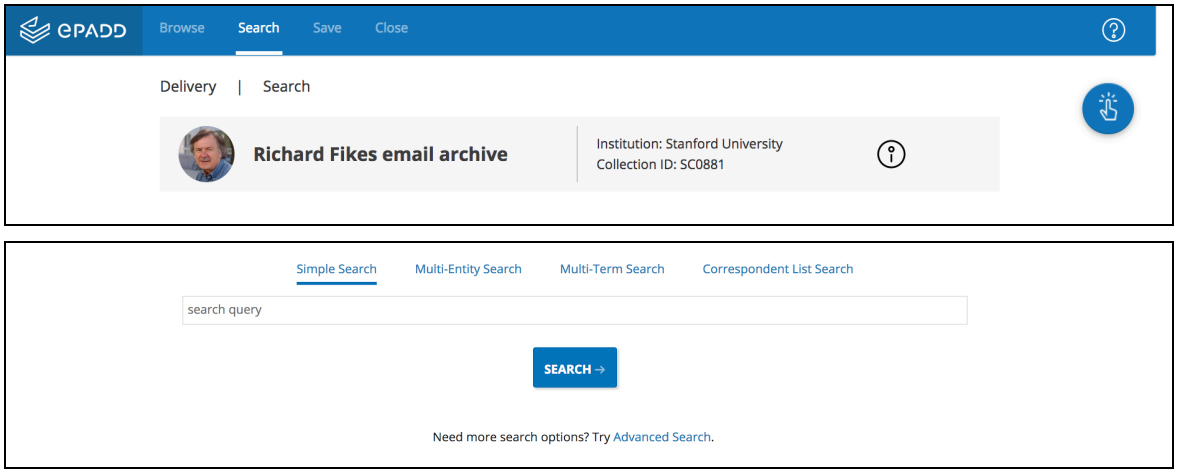
6.3 Browsing the Email Archive

The Browse Menu provides a user with access to several options to browse the email archive, including by Correspondent or by extracted entity type (persons, organizations, and locations). The user can also generate complex tiered searches using a customizable lexicon, as well as browse image attachments, and download documents and other attachments to the workstation for further review.



Browse Menu in the Delivery Module

6.4 Searching the Email Archive



To search the email archive, navigate to the Search screen.

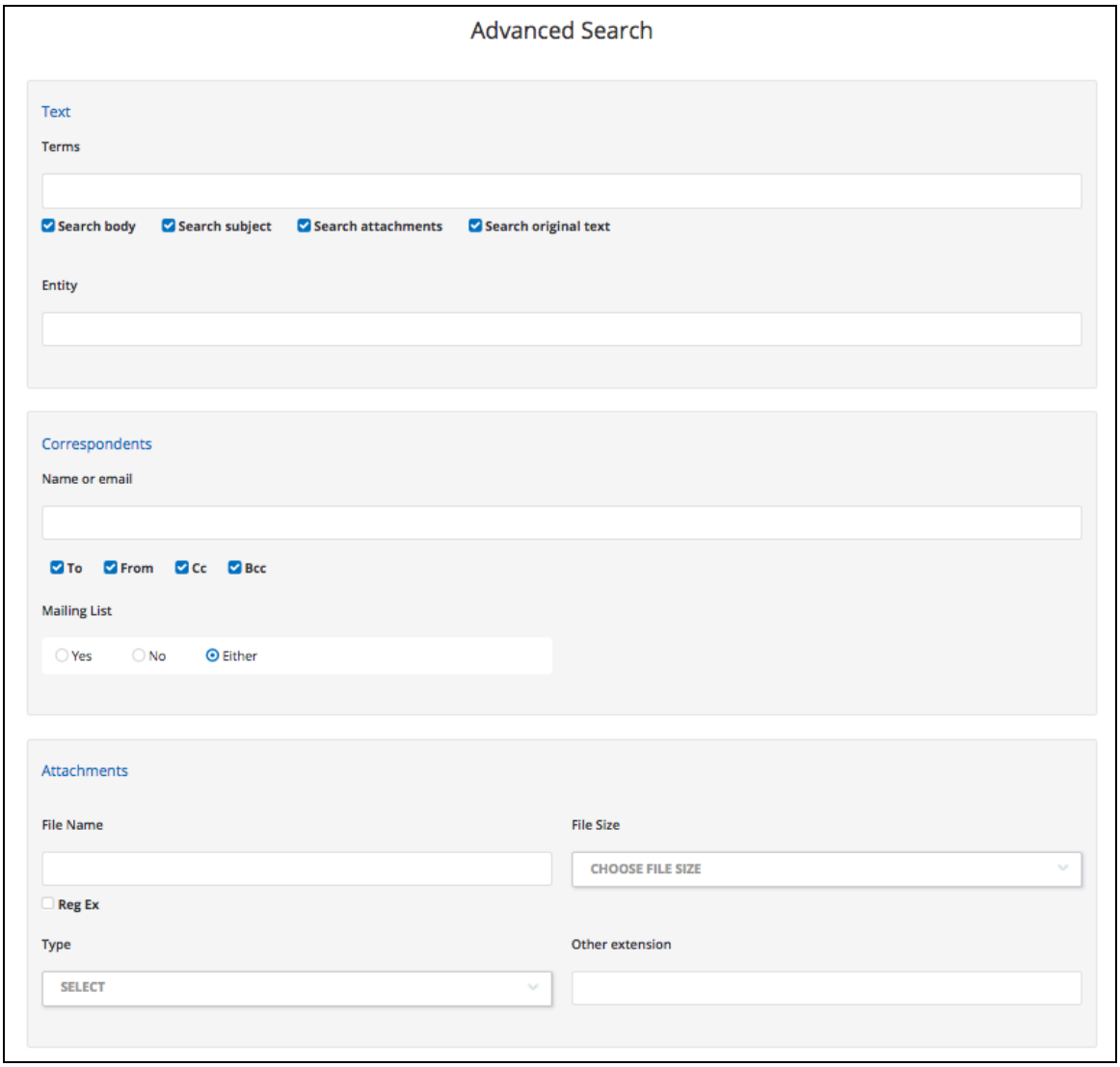
Besides lexicons, accessible via the Browse page, ePADD supports five types of search: simple search, advanced search, multi-entity search, multi-term search, and correspondent list search.

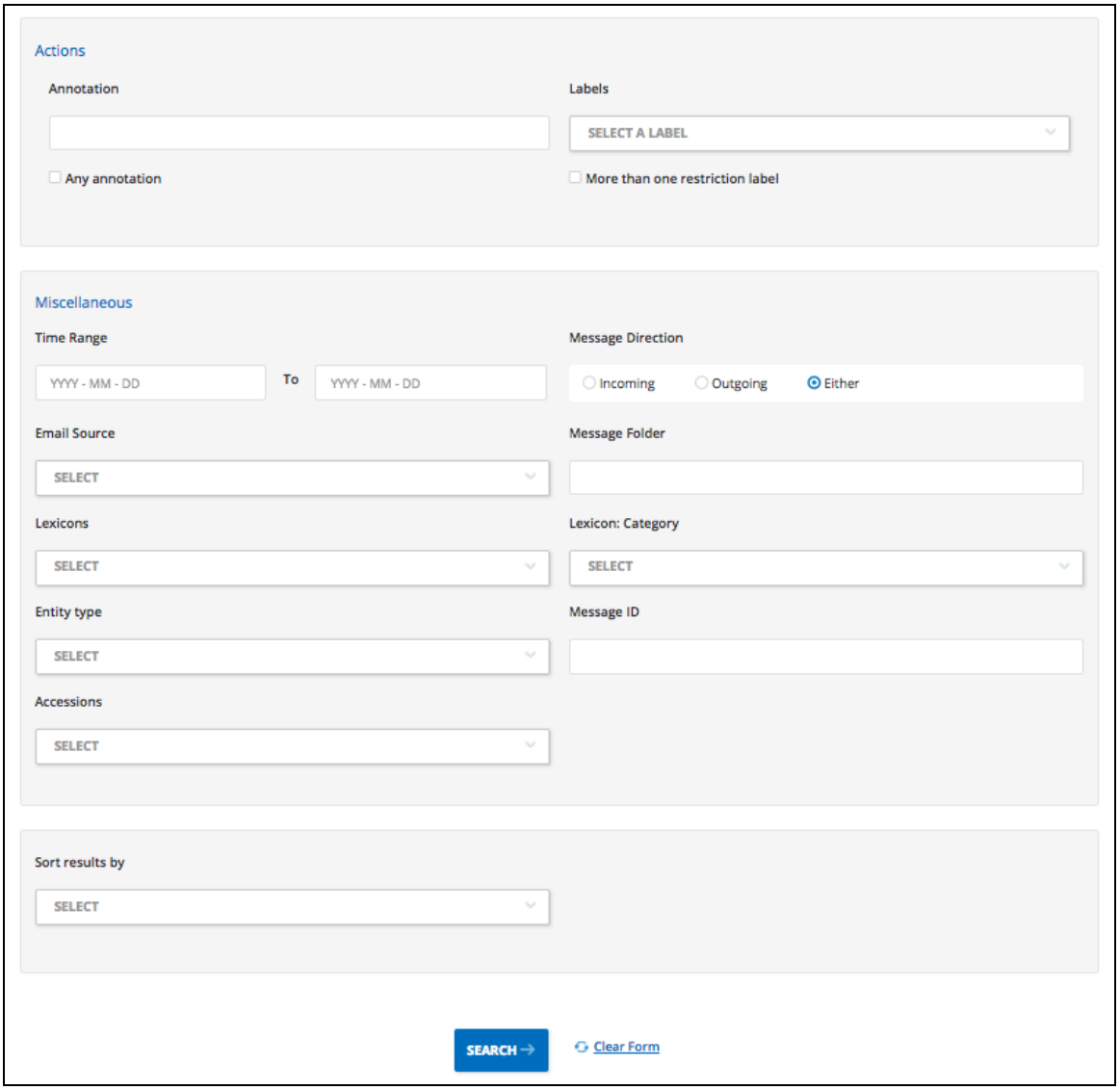
To use the simple search feature, simply type in your query into the simple search box available on the initial search screen. This will search across message headers (correspondents and subjects) and message text, as well as the text of all document attachments.

Once you have entered your term(s) and made your selection, click the *Search* button. The search string will appear highlighted on the message browse screen.

Advanced Search enables the user to to generate complex searches of email text, correspondents, attachments, and any actions previously taken on the messages, as well as a variety of other parameters.

ePADD


Advanced Search
Text
Terms
Search body
Search subject
Search attachments
Search original text
Entity
Correspondents
Name or email
To
From
Cc
Bcc
Mailing List
Yes
No
Either
Attachments
File Name
File Size
CHOOSE FILE SIZE
Reg Ex
Type
Other extension
SELECT



Search Tips can be accessed by selecting the help button in the top right corner of the *Advanced Search* screen.

ePADD also allows a user to search the email archive for entities contained in a given block of provided text, using *Multi-entity Search*. Any matching terms will be highlighted on the results page, and hyperlinked to corresponding messages.

Likewise, ePADD enables a user to search an email archive for multiple terms using *Multi-term Search*. Type or paste one term per line. Any matching terms will be highlighted on the results page, and hyperlinked to corresponding messages.
Finally, ePADD enables a user to search for a list of email addresses using Correspondent List Search. Type or paste one term per line to search for and apply labels to a set of email addresses.

6.5 Labeling messages

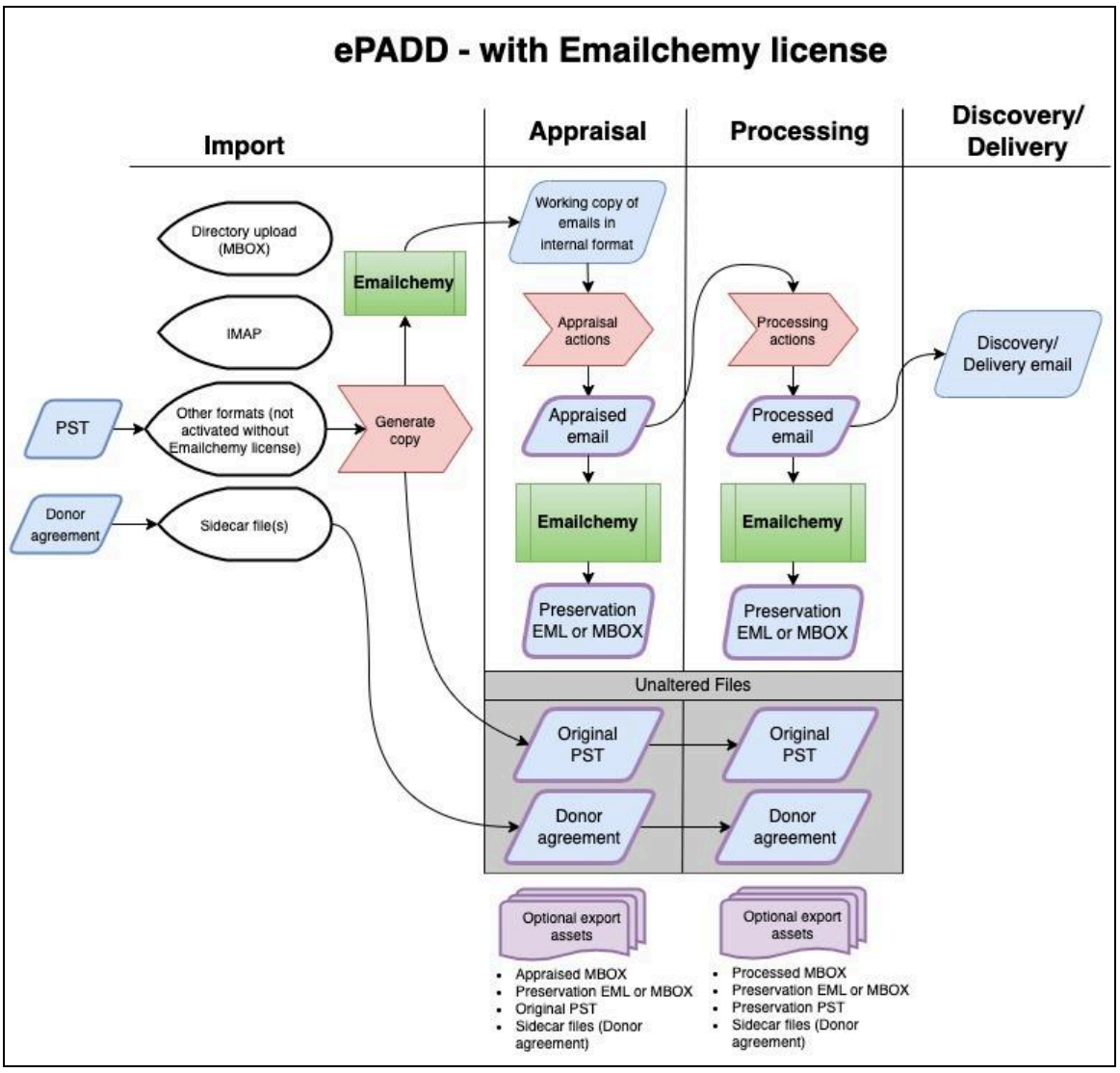
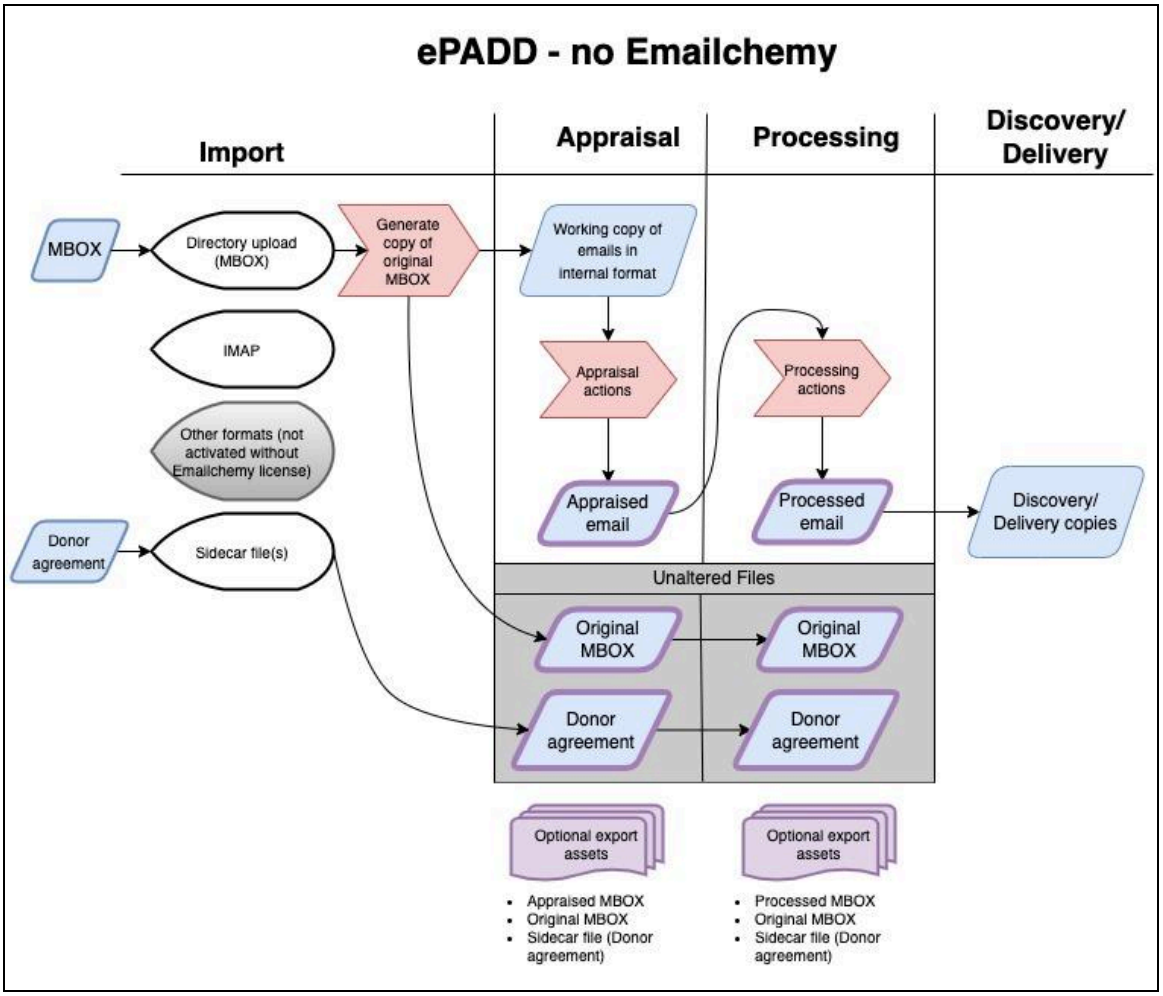
ePADD allows the user to label messages as 'reviewed' to assist with keeping track of research, to add an unlimited number of custom descriptive labels, as well as to request copies of messages from the email archive. Exported messages retain user annotations.

Users can optionally apply labels to a set of messages via the label management interface. Select the label management button, located to the above right of the message pane in the message browse/search result view:

Appendix ePADD Data Flows

The following diagrams illustrate the progression of an email archive in ePADD, from import through to the Appraisal, Processing and Discovery/Delivery modules.

This diagram shows how the email archive moves through the different stages when the optional Emailchemy license is activated:



While this diagram shows how the email archive moves through the different stages when the Emailchemy license is not activated: