Windows MBox Viewer
User Manual 1.0.3.2

Table of Contents

1 Modification History

Revision	Date	Comments
1.0	February 28, 2018	Initial Release
< 1.0.3.1		See ReadMe.markdown 1.0.3.1 included with the release package.
>= 1.0.3.1	April 29, 2019	See the section "What is new and what changed in version 1.0.3.x".

2 What is new and what changed in version 1.0.3.2

2.1 What is new

1. Improved handling of missing character set in the raw email that could result in improper text display.
2. Improved handling of Mail Retrieval Status Text in the Status Bar to reliably display "Mail Retrieval In Progress .." Status text.

2.2 What changed

1. Changed directory structure for printing to files and archiving mails to files. See 6.19.7 section.

3 Feedback

To help to improve MBox Viewer, please post reviews what works and what doesn't, create bug reports and request for enhancements. Provide as many details as possible, such as country, screenshots, etc.

4 Overview

MBox Viewer (or mboxview) is Windows GUI based program to view mbox files such as Thunderbird Archives, Google mail archives or simple Eml files.

The mbox viewer is a simple but quite powerful email viewer that supports the following features:

1. large file support > 4Gb
2. fast parsing of mbox archives
3. quick access to all attachments
4. preview picture attachments
5. zoom in/out a picture
6. export of single mail in Eml
7. print all or multiple selected mails to CSV or Text or HTML or PDF files or send to PDF printer
8. print single mail to Text or HTML or PDF file or send to PDF printer
9. group all related mails as conversations
10. print mail conversation/related mails group to CSV or Text or HTML or PDF file or send to PDF printer
11. open single or multiple selected mails or group of related mails in a browser
12. find user defined text in mail's message and highlight all occurrences
13. search the date, subject, sender, cc, bcc, message text and text appendix
14. sort by date, from, to, subject, size and conversation groups
15. support for mail's header fields and body encoded with different character sets
16. support for mail list editing capability to enable users to compose mail list by merging results of multiple searches and/or by merging one or more user selected mails.
17. support to set Message Window position to Bottom, Right or Left
18. ability to merge/concatenate multiple archive files and remove duplicate mails.

5 Installation

The executable and the source code can be downloaded from Sourceforge or Github.

https://sourceforge.net/projects/mbox-viewer/files/

https://github.com/eneam/mboxviewer/releases

Download the executable package, the latest mbox-viewer.exe-v1.0.3.1.zip, to the working directory and unzip. This will create the mbox-viewer.exe-v1.0.3.1 directory and the stand-alone

mboxview.exe executable, User Guide and scripts sub-directory in that directory. No administrator privileges are required.

6 Running MBox viewer

The mboxview program can be executed by double clicking the executable from the explorer window or from the command line by typing the command name and argument list to it.

6.1 Argument List Summary

The mboxview accept the following command line options.

-FOLDER=folderPath – the full path to the folder containing one or more mbox or eml files. Persists in the registry across multiple runs.

-MAIL_FILE=fileName – full file path or just file name of mbox/eml file to load. If the full path is not specified, the full path is created from the folderPath and the fileName. Doesn't persist in the registry across multiple runs.

-EXPORT_EML=y|n – enables or disables automatic export of eml files generated from mbox archive. Improves traversing performance of mail if set to 'n'. Persists in the registry across multiple runs.

-PROGRESS_BAR_DELAY=seconds – search progress bar is activated when the search duration lasts longer than the value set for PROGRESS_BAR_DELAY. If set to **-1**, the progress bar is disables. Default value is set to **1** seconds. Persists in the registry across multiple runs.

If desired, instead of typing directly desired command line options, a simple .cmd or .bat file can be created for convenience and run from the explorer.

6.2 Setting Options from GUI

The **EXPORT_EML** and **PROGRESS_BAR_DELAY** values can be changed on the fly at any time from GUI.

Click on the "File" to open the drop menu and then select the "Options" option. The below dialog will be presented. Set the desired values followed by OK button.

Note that if exporting of EML files is disabled, you can still export the eml file for the selected mail by selecting "View EML" under the "View" drop menu.

Two options, "**Show Charsets**" and "**Set Character Sets for Header Fields**", can help to properly display header fields in the Message Window. See 6.17.1.1 for additional details.

The "**Picture Viewer**" check box enables/disables the dialog option to preview the picture attachments, see 6.11 for additional details.

The "**Enhanced Select Mail Folder Dialog**" is different version of the "**Select Folder**" option and allows user to see if the mail archives are in the folder. User needs to select any mail file to select

containing folder.

The "**Time**" option allows to display time as Local or GMT.
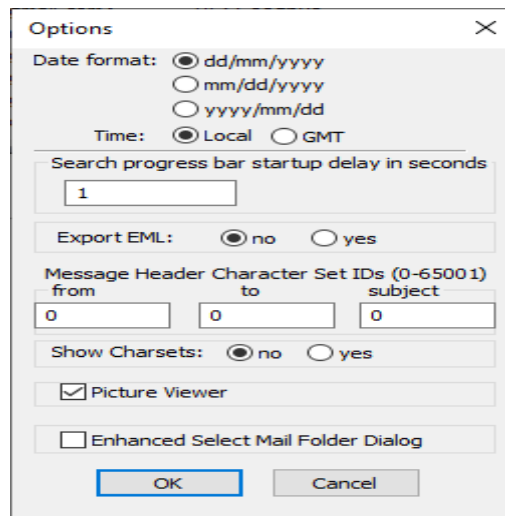
6.3 Basic Use Case

Double click of the mboxview to start the viewer. The picture below shows the screen created by the mboxview after few steps performed by a user. Mail Archive, Summary and Message windows are initially blank after start-up. Steps to load a mail archive and display the content are described next.
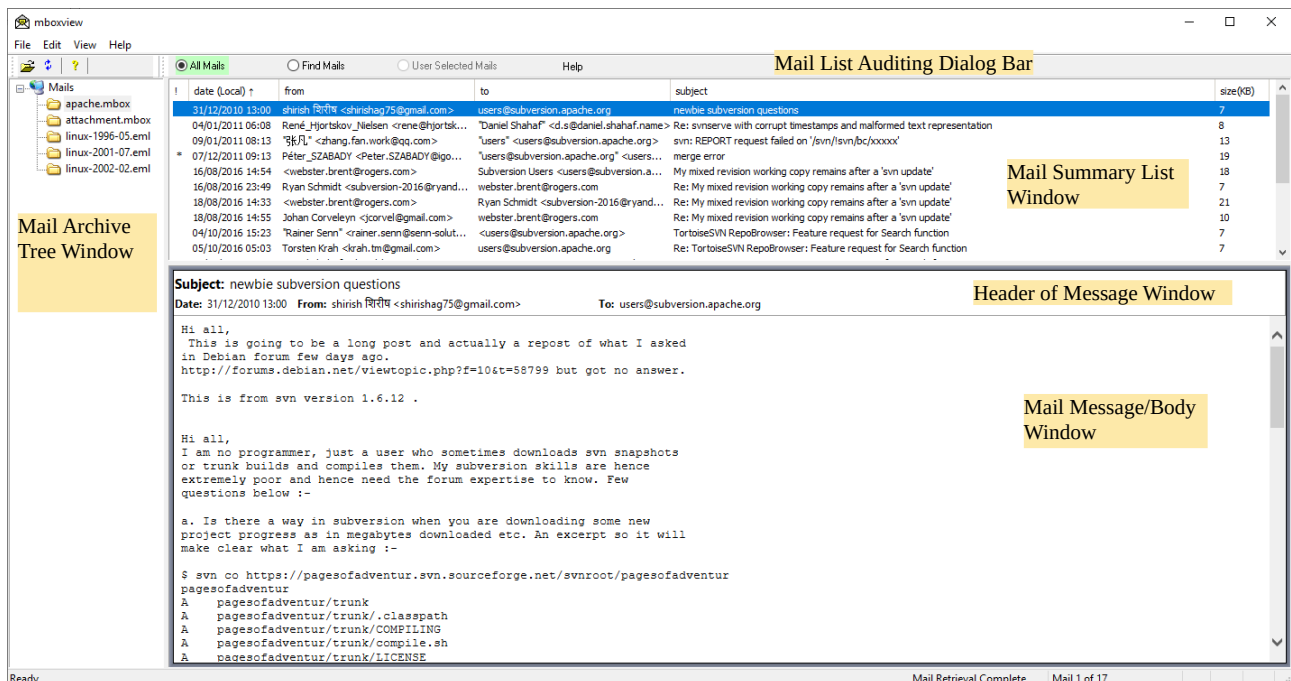
1. Click on the "File" to open the drop menu and then select the "Select folder..." option. Browse to the folder containing one or more mbox and/or eml mail archive files and select it. All valid mail archive files will appear in the Mail Archive Tree window.

Note that different "Select Folder" dialog can be configured via File → Options menu dialog.
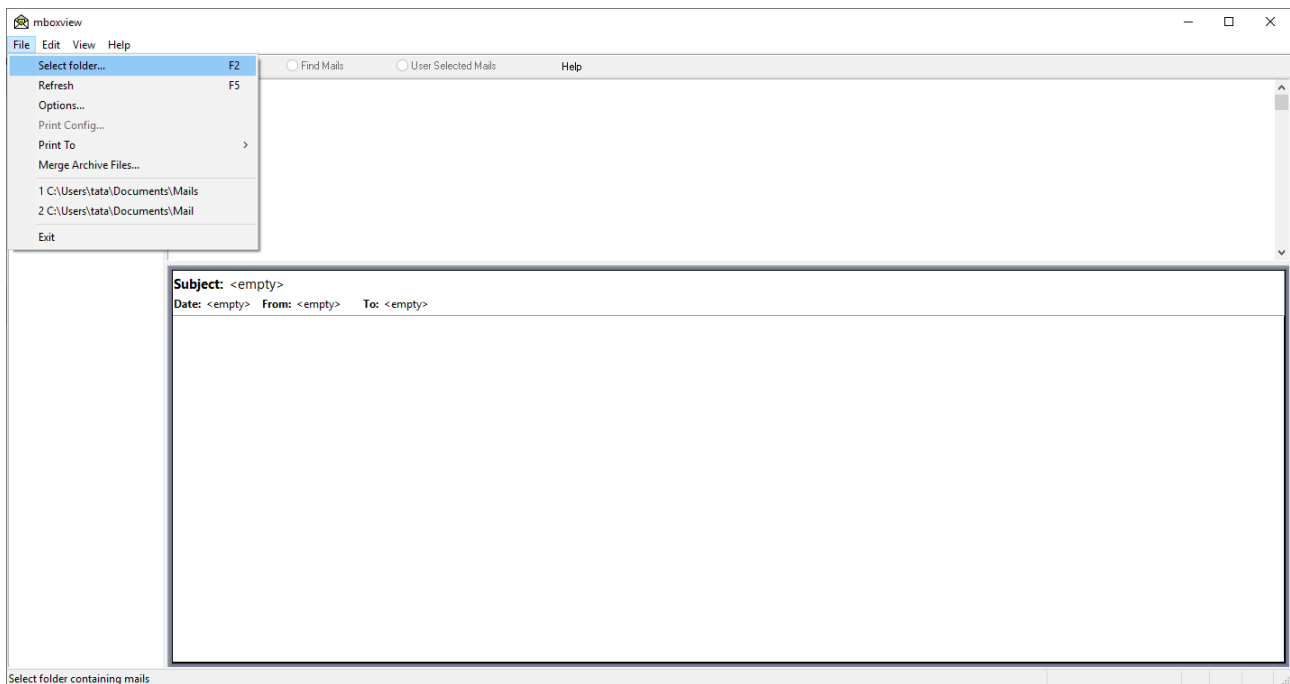
2. Left click on one of the archive files to load all mails within that archive. Progress bar will appear and automatically close after the selected archive is fully processed. Mail header information of each email will appear in the Summary window. Note that parsing of very large archive file may take some time since the mail archives are text files and every character has to be examine one by one.

   However, subsequent loading of mails is done from the index file created by the mboxview during the initial parsing of the archive file and is much faster. The created index file contains content meta data of each mail in the archive file. i.e. the mail header information and the position of each mail within the mail file for quick access to the mail message/body. The index files have the .mboxview suffix.
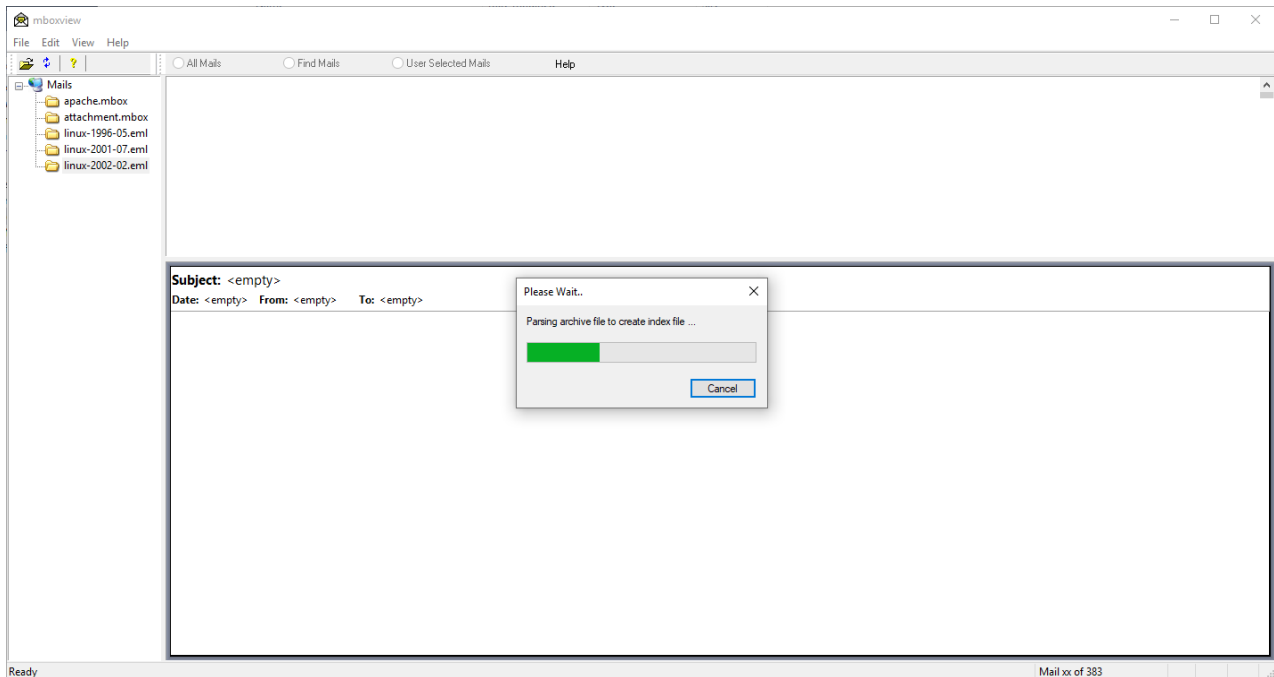
3. Left Click on one of the mails in the Summary window to show the Message/Body of that email in the Message window. The mail retrieval state, total number of mails in the archive and the position of the selected mail within the archive is shown on the status bar. Opening a heperlinks within an external browser is recommended since running Java or scripts or ActiveX controls is disabled in the internal browser to improve security.
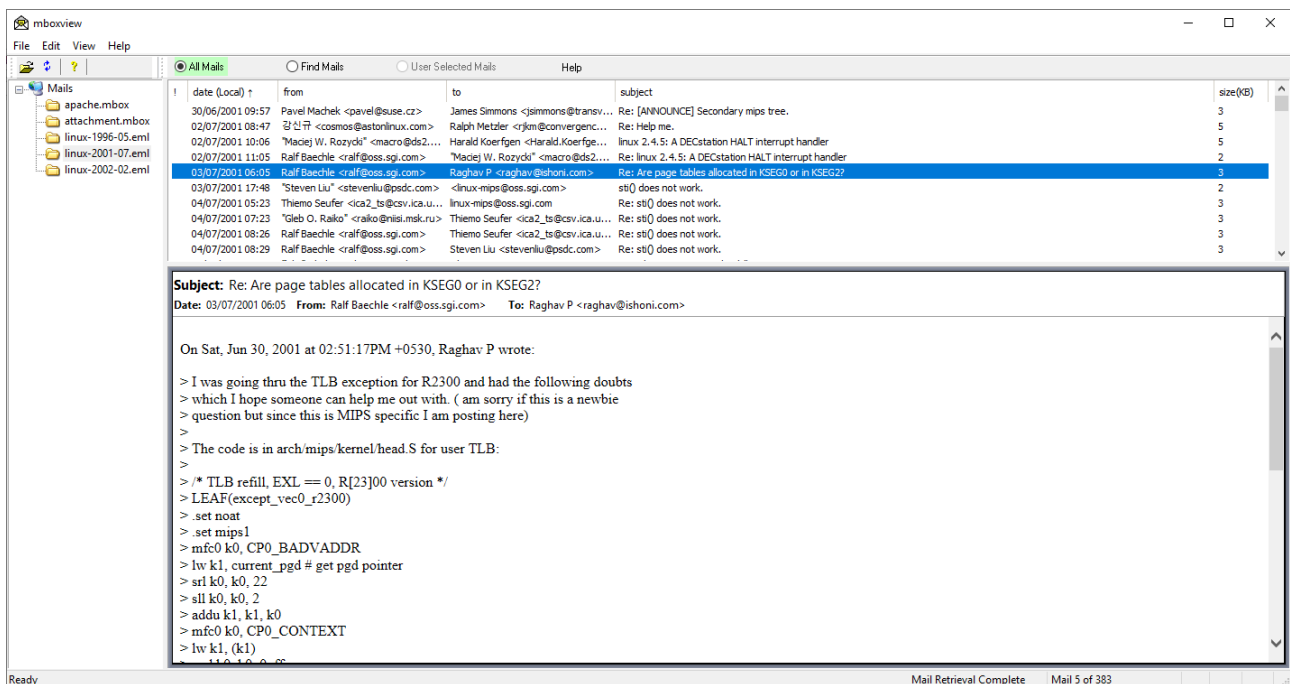
6.4 Message Window Position

Message Window position can be set to Bottom, Right or Left via View → Message Window option. After setting new position, mbox viewer needs to be restarted for the change to take effect.

When position is set to Right or Left, the mail address format will change to show mail name only or mail address if the name is missing.

By default Message Window position is set to Bottom.
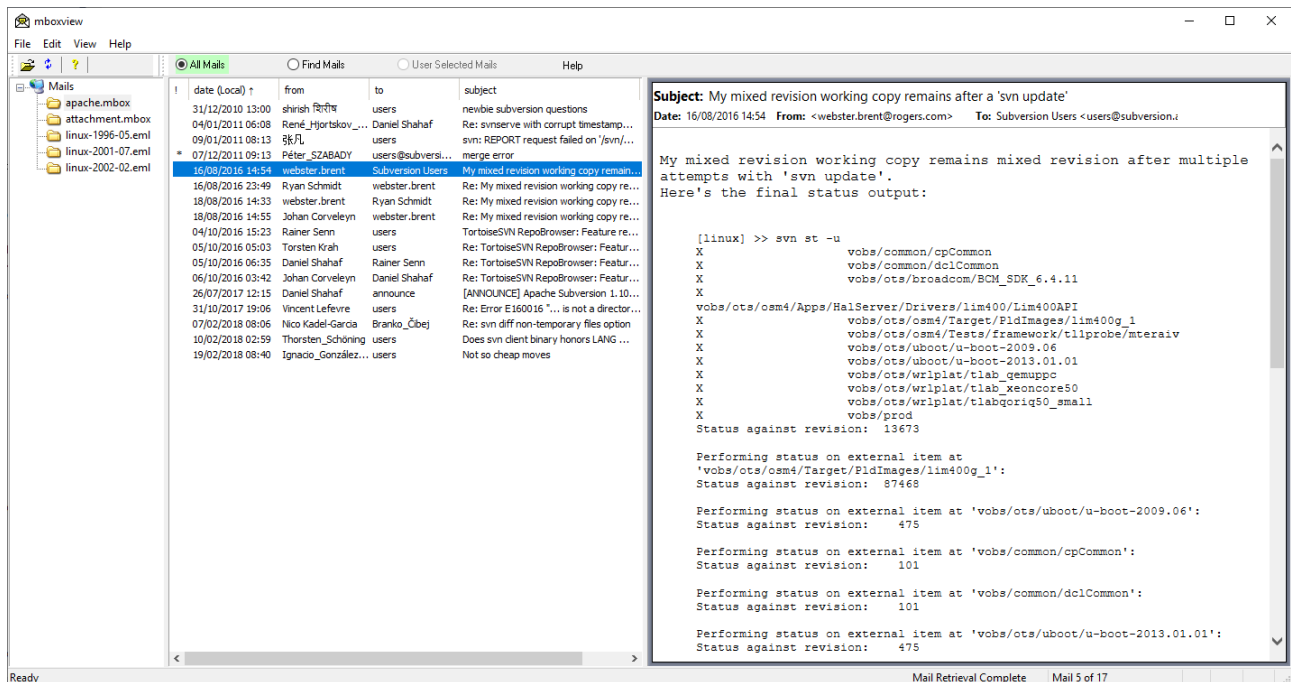
6.5 Single Mail Context Menu

Right Click on a selected mail in the Summary Window to open the context menu as shown below.

The "Print to" context menu allows to print the selected mail to Text, HTML or PDF file or to send to the user selected PDF printer.

The "Open in Browser" menu option will open the selected mail in the external default Web Browser.

The "Find Text" option allows to find user specified text and highlight all occurrences in the Message window.

The "Find Advanced" options will start the advanced search dialog and pre-populate From, To and Subject fields from the mail.
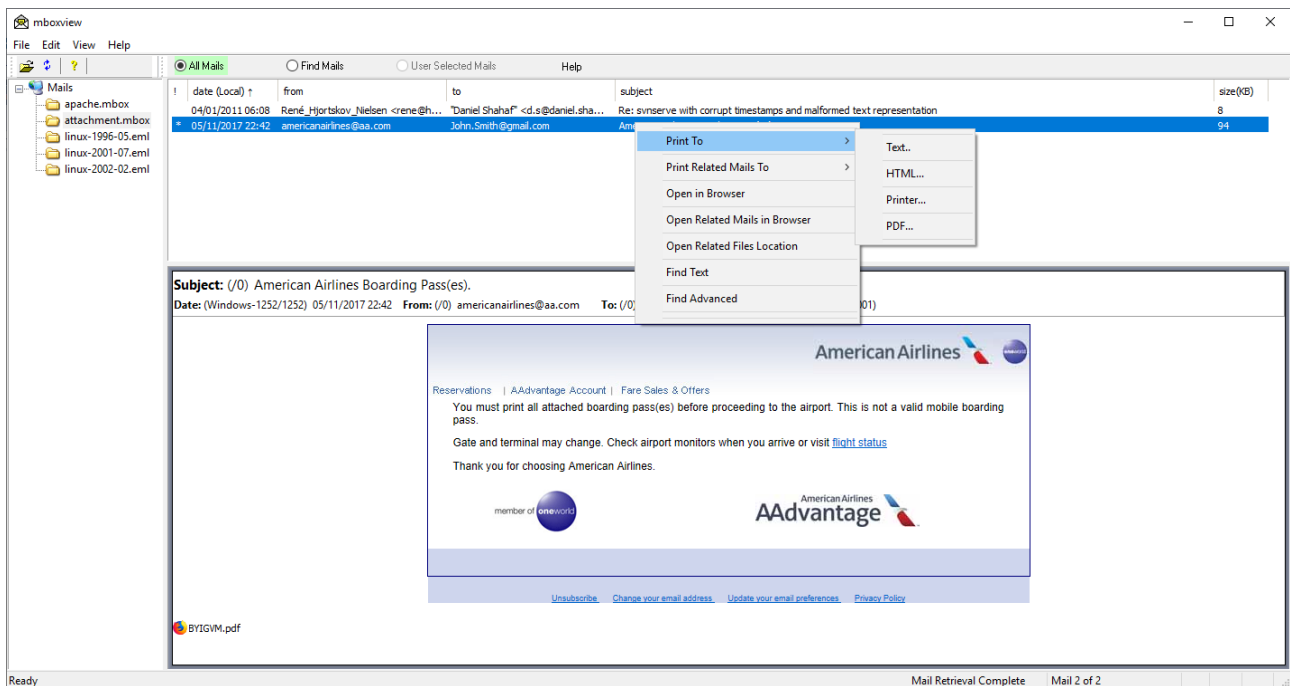
6.5.1   File Names

User can configure file name pattern to apply when printing single mail by selecting File -> Print Config dialog. Pattern is composed from Date, Time, From, To, Subject and Unique Id fields. Only alpha numeric characters are selected from the Subject, other characters will be replaced with the underscore.

Files will be saved in the FolderHousingMailArchives/PrintCache/MailArchiveName folder, where MailArchiveName denotes the archive file name without the .mbox extension.

User must enter the target file name manually when printing to PDF printer, i.e. when selecting Print to → Printer mail option or when printing mail to pdf file from the browser.

Print Config dialog allows users to configure the maximum length of created file name. The maximum file path length supported by Windows is 260 characters for non Unicode applications. Therefore the file name will be limited to (260 – length of FolderHousingMailArchives/PrintCache/ MailArchiveName/MailFileName) path. User can set wanted file name length limit and the mbox viewer will adjust the configured length if needed.

6.6 Multiple Selected Mails Context Menu

Multiple mails can be selected using Windows standard method, i.e. Left click followed by Shift+LeftClick or CTRL+LeftClick or CTRL+A.

Right Click on a multiple selected email in the Summary Window to open the context menu as shown below.

The "Print Selected Mails to" context menu option allows to print the selected mails to single Text or HTML or CSV or PDF files or send to the user selected PDF printer. See also 6.19.6 "Mail

Printing to PDF files by Power Users" section.

The "Open Selected Mails in Browser" menu option will open the selected mails in the external default Web Browser.
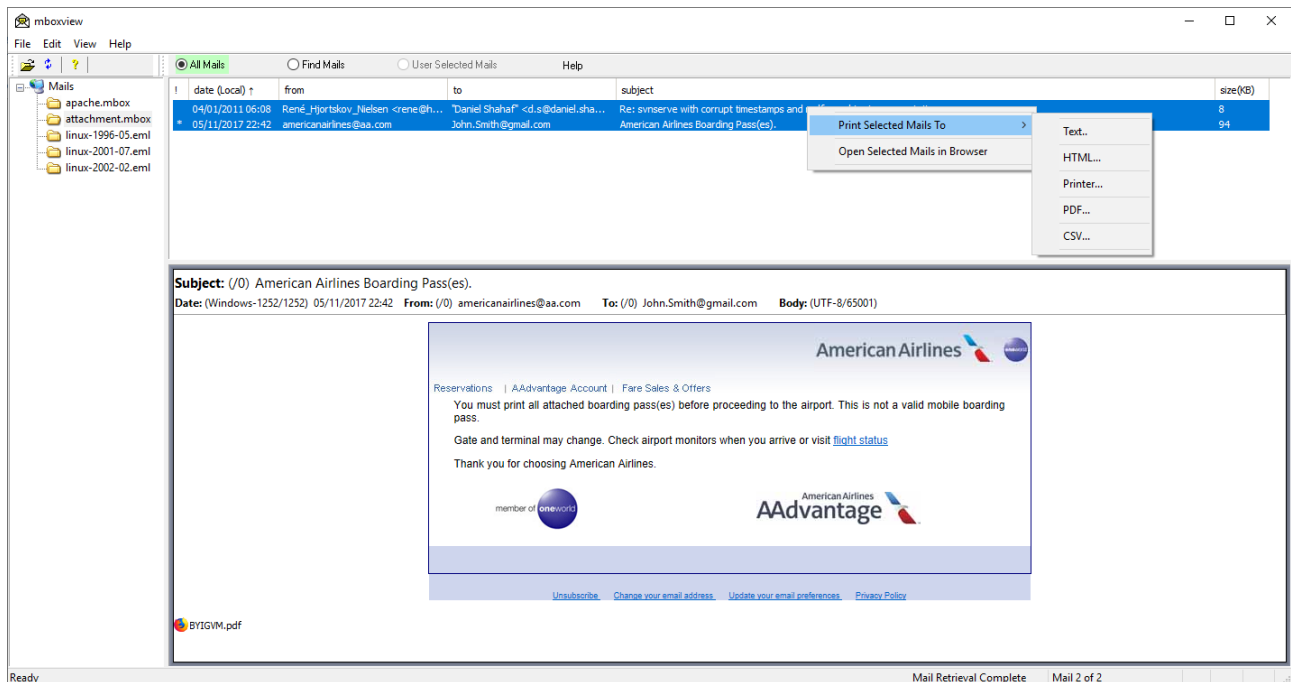
6.6.1   File Names

User can't configure file name pattern to apply when printing multiple mails to a single file. The file name is automatically composed from the archive file name with appropriate extension, i.e. .txt or .htm or .csv or .pdf.

Files will be saved in the FolderHousingMailArchives/PrintCache/MailArchiveName folder, where MailArchiveName denotes the archive file name without the mbox extension.

User must enter the target file name manually when printing to PDF printer, i.e. when selecting Print Selected mails to → Printer mail option or when printing mail to PDF file from the browser.

6.7 Conversation Mail Group Context Menu

All Mails must be sorted by the conversation first before context options can be processed. See 6.13 section.

Right Click on any mail within the selected conversation group in the Summary Window to open the context menu as shown in 6.5 section.

The "Print Related Mails to" context menu option allows to print the selected group of mails (conversation) to Text, HTML, CSV or PDF file or to send to the user selected PDF printer.

The "Open Related Mails in Browser" menu option will open the selected mail group in the external default Web Browser.

6.7.1   File Names

See 6.6.1 section.

6.8 Mail Archive Context menu

Right click on the active mail archive under the Mail Archive Tree Window to open the context menu as shown below.
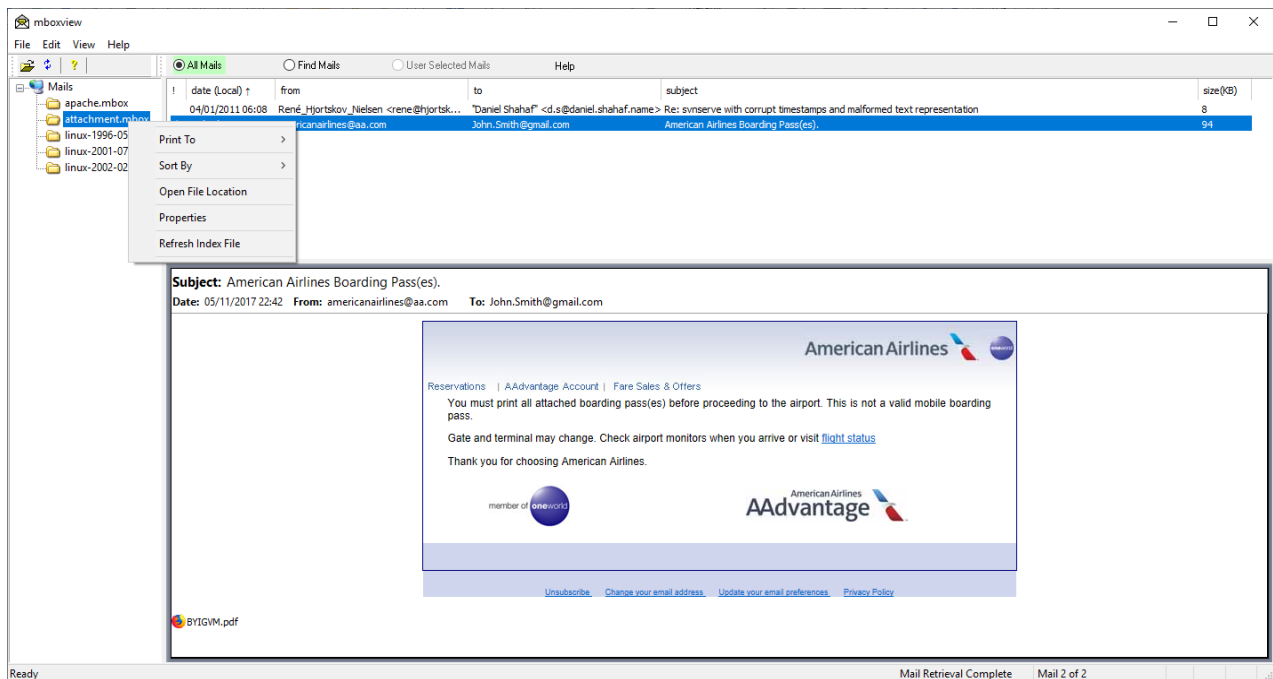
The context menu allows to print all email to single CSV, Text , HTML and PDF files. The name of the created file will be created based on the archive file name with the .csv or .txt or .html or .pdf suffix in the FolderHousingMailArchives/PrintCache/MailArchiveName folder, where MailArchiveName denotes the archive file name without the mbox extension. NOTE that mails will be printed according to the sort order in the summary window.

Size of HTML file can be very large and opening in most of the browsers can be problematic. Microsoft Edge and Google Chrome browser seem to perform best as far as handling large HTML files.

When printing to TEXT file, HTML to Text conversion was added in 1.0.2.8 to handle mails with HTML text block and no Plain text. Conversion can be quite slow and print time could increase significantly compared to 1.0.2.7 in case the large number of conversions is required. Progress bar was added to indicate progress, possibly slow progress.

Printing to single PDF file is the most time consuming. See also 6.19.6 "Mail Printing to PDF files by Power Users" sections for larger scale printing to single PDF file.

All mails of the active archive can be sorted by date, from, to subject, size and conversations in the

Summary Window. This feature is the same as the left click on the selected column in the Summary Window.

The "Properties" option shows the archive file location, size and email count

The "Refresh Index File" option will perform fresh parse of the archive file if needed.

6.8.1　　Print to CSV File Dialog

When printing selected mail archive or selected set of mails by a user to CSV file, user can select what fields to export, the date format and one of two time zones.

Option to normalize content encoding is also available. Default encoding is set to UTF-8 and should not be changed except possibly in special cases.

The "Limit Message to N Characters" value is set by default to 32500 characters and it should not be set higher otherwise Excel and LibreOffice will complain that the character count limit per spreadsheet cell was exceeded. LibreOffice will gracefully handle the text overflow, Excel spreadsheet will get **corrupted**.
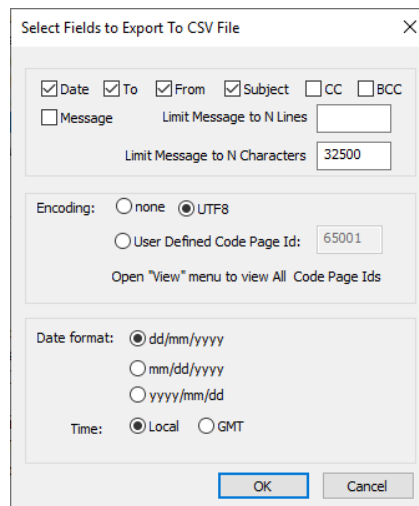
6.9　Mail Attachments

Mail attachments are shown at the bottom of the Message window. Note that the "*' character may appear in the first column for the associated mail to indicate presence of associated attachments. To open location of attachments, double click on the mail in the Summary window or select "View"→"ViewEML" menu option. Temporary folder will open containing all attachments for the email. Use standard Windows applications to view attachments.

You can also double left click to view attachments. Picture attachments will be viewed in the Picture Viewer (if configured in the global Options) and other file types will be viewed by the associated default Windows applications.
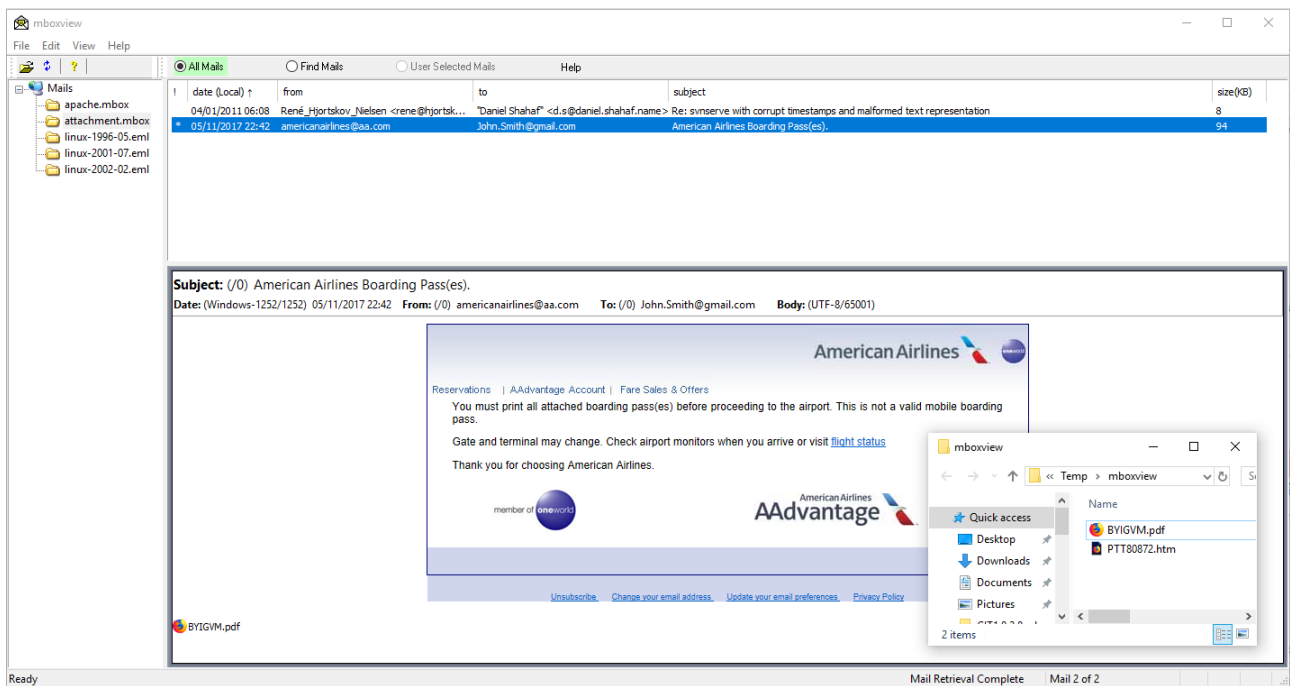
6.9.1   Attachment context menu

Right click on the selected attachment to open the context menu. The context menu allows to open or print the attachment by the associated default application. You can also open the folder housing the attachment. The selected attachment will be highlighted in the folder.
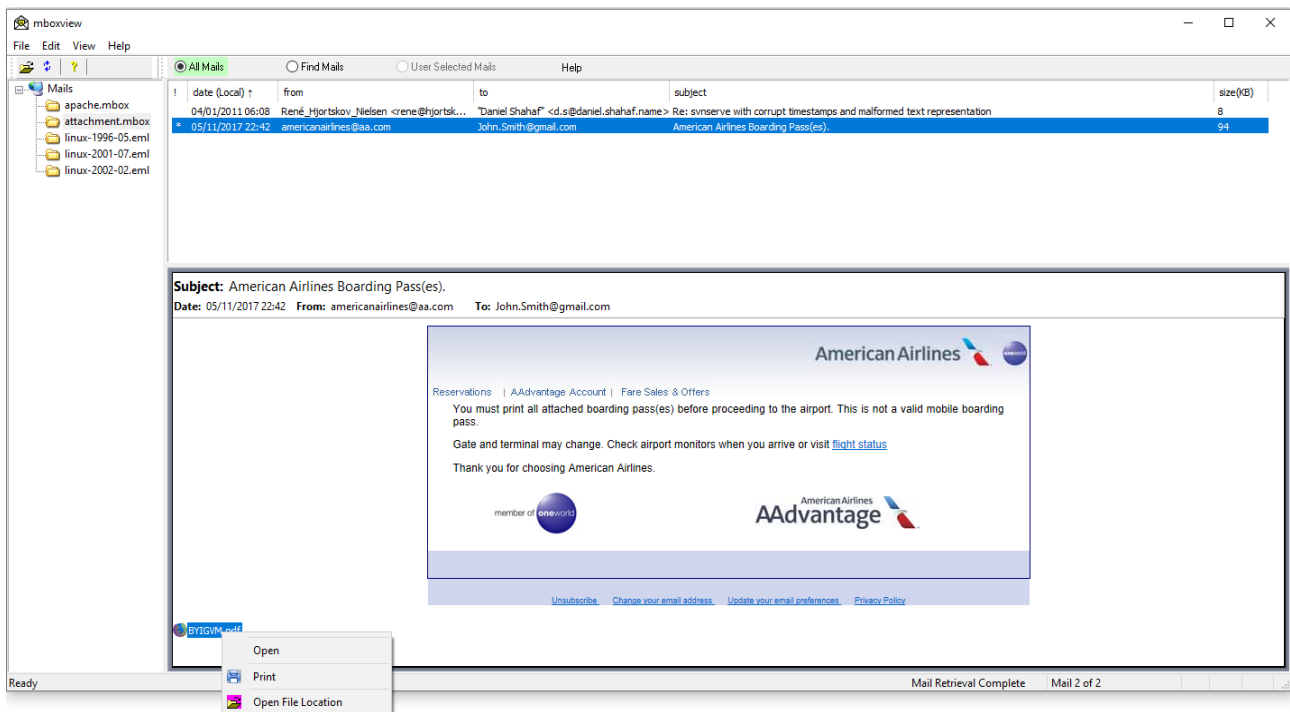
6.10   Inline Image Cache

The mboxview will create the cache of inline images to support images embedded directly into the

mail body/message. The cache is created during the initial parsing of the mail archive file. Separate cache sub-folder is created per each archive file (assuming at least one inline image was found) in the FolderHousingMailArchives/ImageCache folder. If the inline image cache is missing for any reason, try to refresh the index file by selecting "Refresh Index File" context option on the respective file archive.

6.11   Picture Viewer

Double left Click on any attachment to start Picture Viewer dialog, example below.
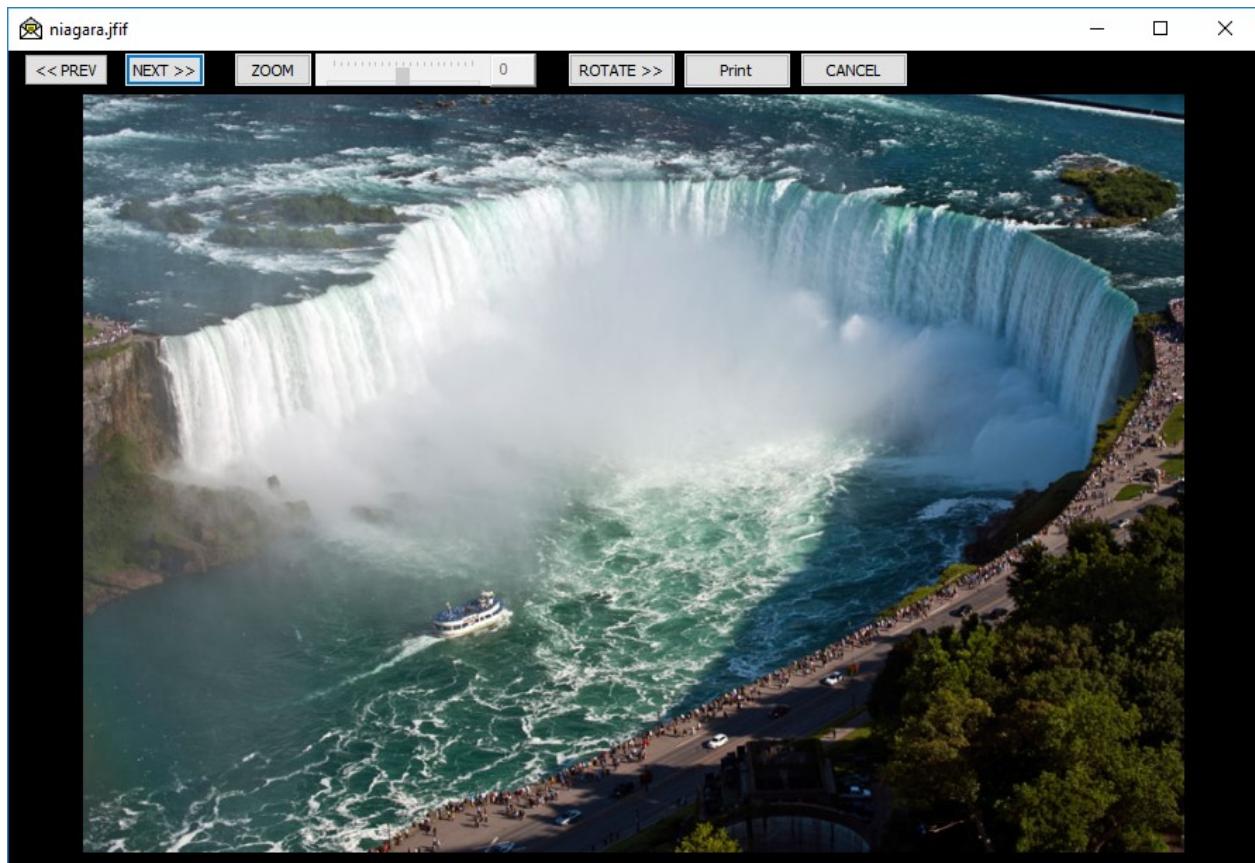
The Picture Viewer supports files with the png, jpg, pjpg, jpeg, pjpeg, jpe, gif, bmp, ico, tif, tiff, jfif, emf, wmf, dib suffixes.

The NEXT and PREV controls enable user to browse all picture/image attachments associated with the selected email. You can select specific attachment to be displayed. You can later browse through all picture attachments.

The ROTATE control allow to rotate clockwise the current/visible picture.

The ZOOM control allows to zoom in/out picture. Click on ZOOM to enable/disable zooming. Slide to right to enlarge an image, slide to left to shrink. Enlarged image can be dragged using the left and scroll buttons of the mouse. Click ZOOM to disable zooming and reset the image to original size.

The PRINT control allow to print via default application.

The Picture Viewer is enabled by default but it can be disabled via FILE→Options dialog. Setting persists in the Registry across multiple runs.

All attachments are also accessible by double clicking on the selected mail in the Summary window or by selecting "View"→"ViewEML" menu option.

6.12   Mail Navigation

Click on a specific mail in the Summary window and use **UP, DOWN** keys to move to the next or previous mail respectively.

Use **HOME** key to jump to the first mail. Use **END** key to jump to the last mail.

Use the mouse wheel or vertical bar to scroll up and down mails and then select one of the mails.

6.13   Mail Sorting

The Summary window is divided into 6 columns: attachment indicator, date, from, to, subject and mail size. Each column can be sorted by clicking on the column title. Subsequent click on the same column title will reverse the sorting rule. By clicking on the first column title, emails are sorted by conversations and mark by different colour as shown below.

Mail can also be sorted via the mail archive context menu or by selecting "File"->"Sort By" menu option.
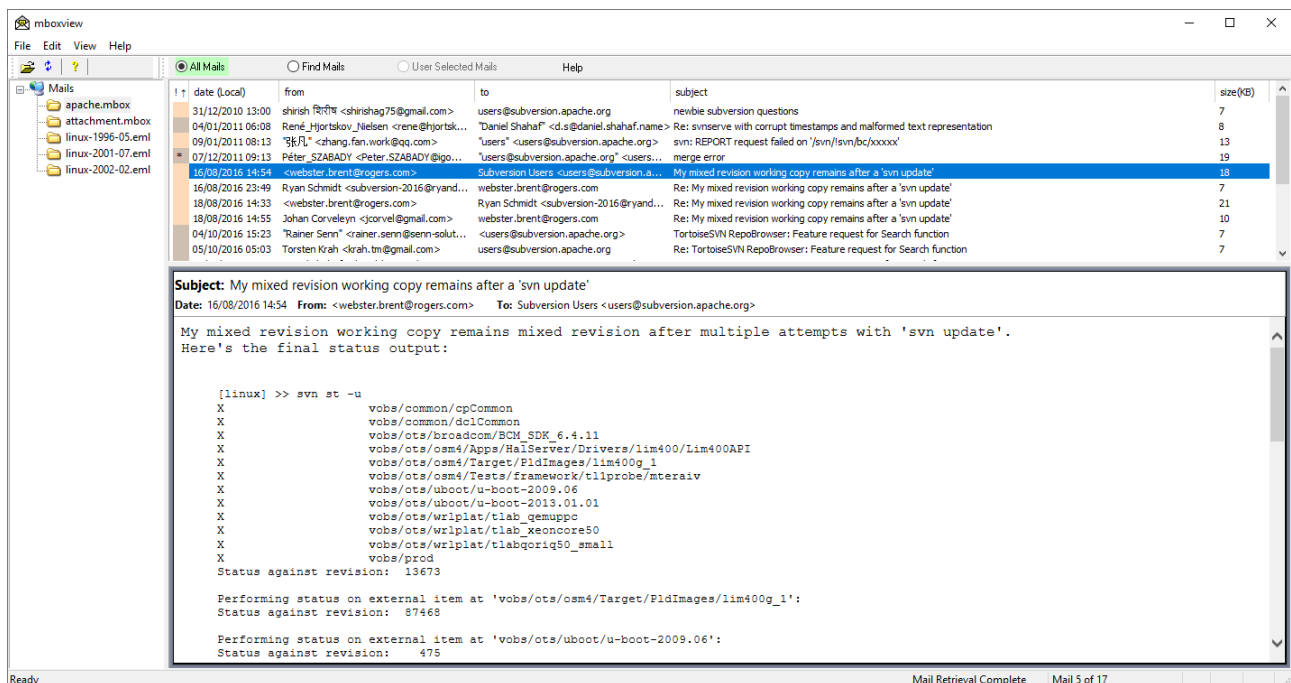
6.14　　　Dynamic Column Resize

Columns in the Summary windows are automatically resized when the window is resized.

However, if columns are resized manually, some of the columns may no longer be present in visible area/window.

You may need to use the scroll bar to see missing columns, otherwise maximize and then restore the window to see all columns within the Summary window again.

6.15　　　Mail Searching

Mbox viewer supports two search options, i.e. basic Find and Advanced Find.

The basic Find option allows users to specify single keyword for searching across header fields, mail text and attachments.

The advanced Find option allows user to specify keyword per each header field, mail text and attachments.

6.15.1　　Basic Find

The basic Find option allows users to specify single keyword/string for searching across header fields, mail text and attachments.

Users can configure the search scope by selecting all or subset of the from, to, subject, message and text attachments options. The search succeeds if the specified keyword matches any of the checked fields.

Click the "Edit" drop menu and select "Find" option to configure and start the search. Alternatively the CTRL+F will also present the search dialog.

The search dialog box allows user to specify the search string, whether to match the whole word, whether the search is case sensitive, configure the search scope, the start and end dates to consider for filtering, the Next/Previous search direction and whether to highlight all occurrences of the matched text.

Note that the Next search start with the oldest email while the Previous search starts with the latest.

The search logic simple, i.e. mail is considered a match if the search string is found in any of the checked fields in that mail.

To repeat the search in order to find the next matching mail, click the "Edit" drop menu and select "Find again" option, or select F3 key.

If "Find again" or F3 key is selected without prior "Find" or CTRL+F, the search dialog box will be presented to a user.

6.15.1.1   Find All Occurrences Option

User can check "Find all occurrences" to find all matching mails. Automatic highlighting of all keyword occurrences is not supported in this case. Matched mails will be displayed under the Found Mails List and the corresponding button in the dialog bar will be highlighted. User can left click on All Mails button to switch mail list back to All Mails.

The basic Find can also be run while in the Found Mails list and while in User Selected Mails list when the list is enabled.

6.15.2 Advanced Find

The advanced Find option allows user to specify different keyword per each header field, mail text and attachments.

Users can configure the search scope by selecting all or subset of the from, to, subject, cc, bcc, message and text attachments options.

Click the "Edit" drop menu and select "Find Advanced" option to configure and start the search.

The Find Advanced is also available as menu option for the selected mail. Right click on the mail and select Find Advanced option. Mbox viewer will pre-populate From, To and Subject fields in this case.

Only checked fields are use by matching logic and search strings are ignored if present unless field is checked.

The search logic is hard-coded for the most common case and is as follow:

(unidirectional or bidirectional and operator of From and To) and Subject and CC and BCC and (Message or Attachments)

More complex searching can be achieved by leveraging User Selected mails list. User can run multiple advanced searches and merge results into User Selected Mails list.

6.15.3 Search Reliability

Searching may not work as expected in all cases due to different character sets on input and in the mails.

Base64 or "quoted-printable" text content blocks and text attachments are first decoded before the search to eliminate false negative results.

Searching of the header fields is performed on the received raw data.

Raw mail data may contain plain text and/or html text blocks. Both encode the same/similar content. If the mail contains plain text blocks, only plain text will be searched. If the mail doesn't contain any plain text blocks, the html text blocks will be searched.

False positive results may happen when searching text/html content blocks since the html tags are not ignored when searching. The search string may match one of the html tags and result in false positive match.

If possible, use ASCII characters only to compose the search string.

6.15.4 Special Search Key

The single character '*', i.e. the star character can be specified as the search string to enable traversing the subset of mails between the start and end dates.

6.15.5 Search Progress Bar

Search progress bar displays the progress along the configured search string and enables user to cancel the search is so desired. The progress bar is not created unless a given search duration lasts longer than the delay time controlled by -PROGRESS_BAR_DELAY command line option or specified via GUI. Default value is 1 second.

6.15.6 Search Performance

Searching of header fields performs best since all fields reside in the memory.

Searching of messages and text attachments may take some time for very large files, say larger than 10-20 Gbytes. Search requires to read the content blocks from the email file and decode the content when required. Worse case scenario is when the specified search string is not present in the mail

archive. When activated, the search progress bar shows the configured search string and enables user to cancel the search is so desired.

Note that when searching of message and text attachments is configured, all mails will be sorted by date first, except when sorted by conversations, to maximize search performance.

6.16 Refreshing Archive Tree Window

If an additional mbox or eml file is placed in the already selected folder, the Archive Tree window needs to be refreshed to see the new additions. This can be done by selecting "Refresh" option under the "File" drop menu or by clicking on the button showed below.
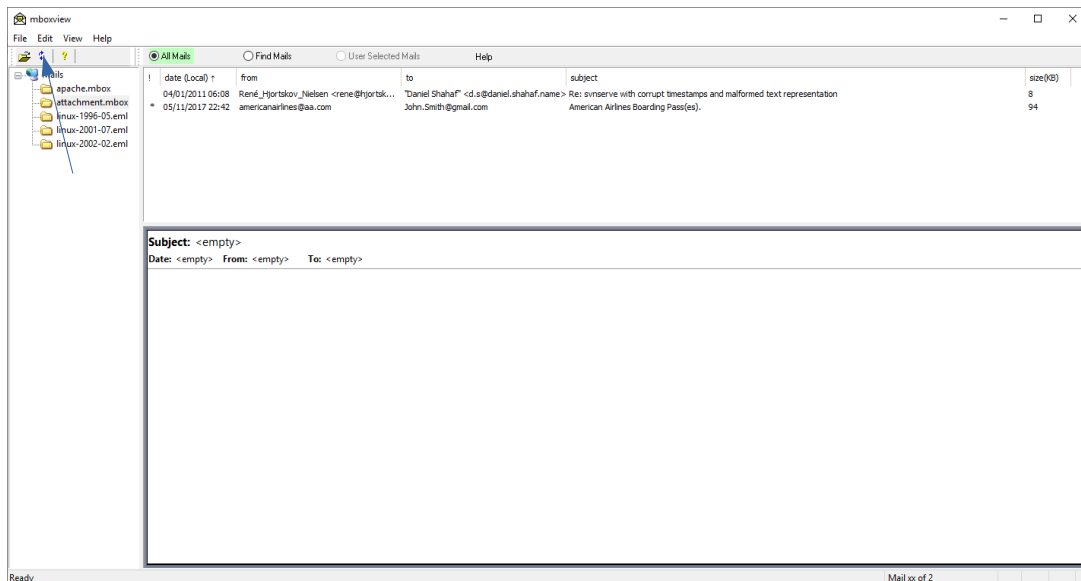
6.17 Language (Character Sets) Support

6.17.1 Message Window

Different languages (character sets) are properly supported in the Message Window. This is thanks to the Microsoft Web Browser emulation/control class integrated into the mboxview. The mboxview generates the htm file from the mail's body and Web Browser displays the generated file in the Message window.

Display may not be correct if the character set for the mail's body is not defined in the mail. To resolve the incorrect display, you can try to apply different encoding by right clicking within the window, selecting "encoding" and "more" options to apply different language.
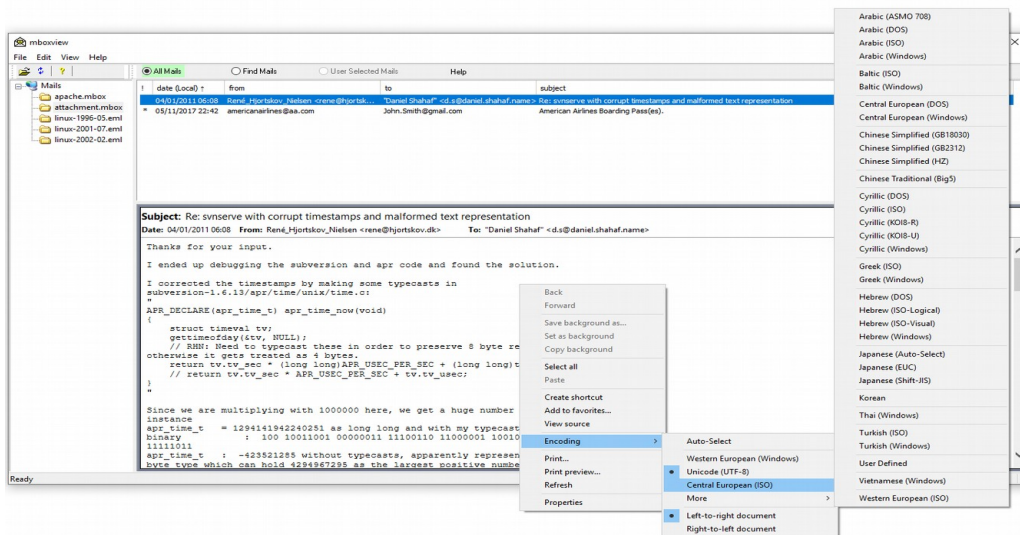
6.17.1.1   Header of Message Window

The text of the mail header fields is displayed correctly as long as the corresponding field encoding type is present in the email.
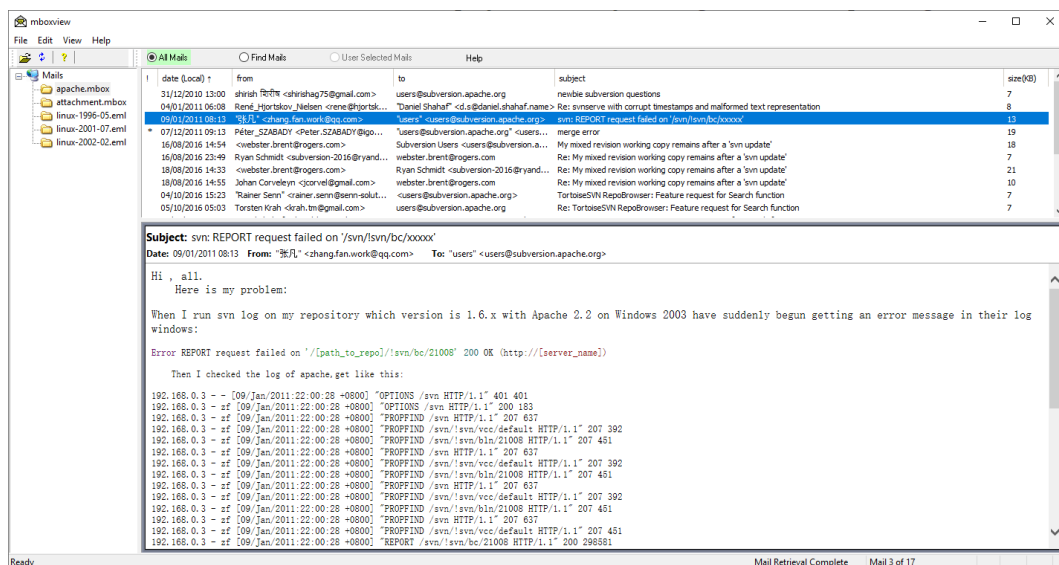
Two global options can be used to address incorrect field presentation in the Message Header in case the field encoding is not present in the email. Missing field encoding type may or may not be a problem. It should not be a problem if the default/local character set installed on your computer matches the original encoding of the header fields. It might be an issue if you are receiving emails from different countries.

As shown below, you can see all encoding/character sets applied to the header fields by enabling "Show Charsets" option. To enable, select "Options" from the "File" drop menu and enable "Show Charsets".
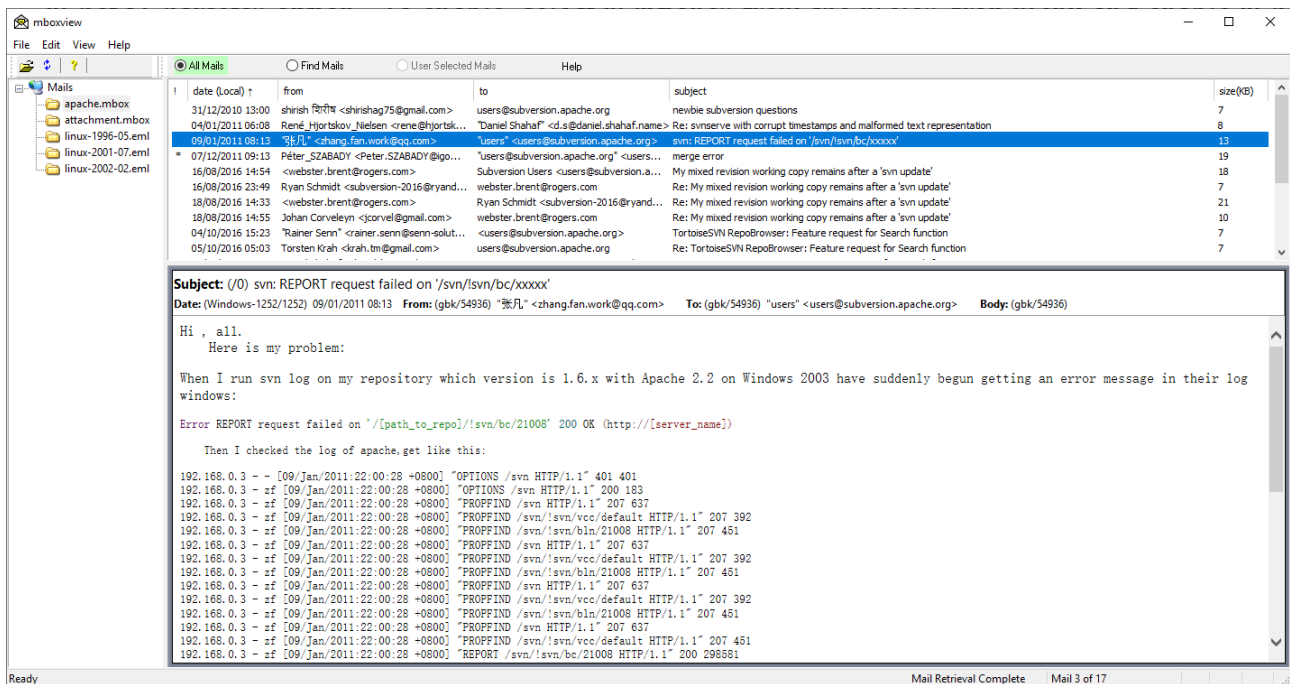
If the field text presentation is not correct, you can try to configure character set (or so called code page according to Microsoft terminology) for the selected header fields and see if that helps. Per field code pages can be configured via Options dialog box. Note that the user provided encoding doesn't override encoding in the email, if present, which is denoted by the code page greater than zero. Code Page set to zero indicates missing or unknown character set.

If you change the code page value(s) in the Options Dialog box, you need to refresh the selected email by going up and down (or down and up).

All Windows supported code pages are listed in chapter 7 or can be viewed in the external Browser by selecting "View"→"View Code Page Ids".

The character set information appears after the field title and is formatted as "(character set name/code page)". The '*' star character is appended after the code page if the user specified code is applied). Example "(UTF-8/65001*)".

The character set information after the "Date:' title shows the local/default character set installed on your computer.

6.17.2   Summary Window

Mail header fields are displayed correctly as long as the corresponding field encoding type is present in the email.

6.18   Mail List Auditing

Mail list auditing capability enables users to compose mail list from all mails by merging results of multiple searches and/or by merging one or more user selected mails.

To support list auditing, the Mbox Viewer maintains 3 internal mail lists:

- **All Mails** list is populated from the selected archive file under the Mail Tree.
- **Found Mails** list is populated by the search results. User can run Find Advanced dialog or set the Find All Occurrences option in the Find dialog.
- **User Selected Mails** list is composed by a user from the mails in the All Mails and Found Mails lists. Mails on All Mails and Found Mails lists are marked by the red vertical bar in the first column if they are also on User Selected Mails list.

Each internal mail list has associated button in the dialog bar located next to the tool bar. When a particular mail list is shown in the Mail Summary Window, associated button is highlighted.

Access to the User Selected Mails list is disabled upon startup and the associated Button gray out. It can be enabled by the user to perform simple list auditing if desired by selecting View->User Selected Mail List to enable/disable.

When User Selected Mails list is enabled, additional mail menu options will also be enabled such as Copy Selected Mails to User Selected Mails.

Content of the User Selected Mails list is controlled by the user. User can merge search results with the content of the User Selected List.

User can run the search multiples time and merge results multiple times with the User Selected List. Search results can be pruned before merging.

In addition, user has an option to select/highlight one or more mails in the Summary Mail Window and copy to the User Selected Mails list.

Standard Windows method is used to select and highlight multiple mails, i.e. Shift+Left Click and CTRL+Left Click and CTRL+A.

Merging/copying process will not create duplicate mails in the User Selected List.

User can select and highlight one or more mails in the Summary Mail Window and remove from the active list. Mails can't be removed from the All Mails list.

All Mails list content persists until new mail archive is selected.

Found Mails list content persists until new search or when new mail archive is selected.

User Selected Mails list content persists until cleared by the user or new mail archive is selected.

6.18.1   Mail List Archiving

Found Mails list and User Selected Mails list content can be saved to new mbox archive files or just to the .mboxlist mail list files.

Mail Archive and Mail List files will be created in the FolderHousingMailArchives/ArchiveCache/ MailArchiveName folder.

Archive file created from the Found Mails list will be created by appending _FIND suffix to the base name of the main archive file name.

The _USER suffix will be appended when creating archive file from the User Selected Mails list.

For example, if the main mbox archive file from gmail is called "All mail Including Spam and Trash.11.09.2018.mbox", then created archive file will be named "All mail Including Spam and Trash.11.09.2018_USER.mbox."

Separately, user can save mails in User Select Mails list to .mboxlist file, for example "All mail Including Spam and Trash.11.09.2018_USER.mbox.mboxlist". The .mboxlist files are much smaller than the mail archive files. To protect against catastrophic failure of the mbox viewer while saving mails to .mboxlist file, the mbox viewer maintains "All mail Including Spam and Trash.11.09.2018_USER.mbox.mboxlist.bak1" and "All mail Including Spam and Trash.11.09.2018_USER.mbox.mboxlist.bak2" files. To recover, user must copy one of these two files to "All mail Including Spam and Trash.11.09.2018_USER.mbox.mboxlist" file.

User can reload the last mail list file into the User Select Mails. Reload can be requested at any time including after restart of the mbox viewer.

After mbox viewer is restarted or when User Selected Mails List is empty, User Selected Mails list can be restored when All Mails list is active.

When User Selected Mails list is active and not empty, it can be reloaded by a user from the last .mboxlist file.

WARNING: The .mboxlist file will no longer be valid if the master mail archive is changed.

Merging Multiple Mail Archives
Multiple archives can be concatenated into a single archive using "File -> Merge Archive Files" option. This is useful if you need to analyse multiple mail views, provided as separate archive files, derived from the same archive file.
After merging is completed, select "File -> Refresh" to request the mbox viewer to discover new archive.
In order to remove duplicate mails from the concatenated file:

1. Select the created archive.
2. Select "View -> User Selected Mails" to enable User Select Mails list.
3. Right click on any mail to select "Copy All into User Selected Mails" option to copy all mails into User Selected Mails list.
4. Click on the "User Selected Mails" button to select User Select Mails list.
5. Right click on any mail to select "Remove Duplicate Mails" option to remove duplicate mails from User Select Mails list.
6. Mails are considered duplicate if the Date, From, To header fields and unique message ID generated by the mail provider match.
7. Select "Save as Mail Archive file" option to save mails on the User Selected Mails list.
8. Select "File -> Refresh" to discover new no duplicate mails archive file by the mbox viewer.
9. You may want to rename the default name assigned to the archive file before the above step.

6.19   Mail Printing Overview

Mbox Viewer supports printing of all, single or multiple mails to CSV, Text, HTML, PDF files and to PDF printer. Mails can also be printed to PDF from any Web Browser by opening mails printed to HTML files.

By default all, single or multiple mails are printed to single CSV, Text, HTML and PDF files without any additional configuration or manual steps. However there is a limit how many mails can effectively be printed to a single file. Option is provided to print mails to separate files to scale to larger number of mails. See 6.19.6 section.

The next sections below will cover these limitations and alternate solutions.

6.19.1   Mail Printing to CSV Spreasheet file

By default, Mbox Viewer supports printing of all mails or selected group of mails to single CSV file. All mails can be printed to single CSV file but a particular Spreadsheet tool may limit the maximum supported size of the spreadsheet file.

User can select and print groups of mails to separate CVS files as a work around. There is no option to automatically split all mails into configurable number of mail groups for printing to separate CSV files.

6.19.2   Mail Printing to TEXT file

By default Mbox Viewer supports printing of all, single or selected group of mails to single Text file. All mails can be printed to single TEXT file but a particular Text viewer tool may limit the maximum supported size of the text file.

User can select and print groups of mails to separate Text files as a work around. There is no option to automatically split all mails into configurable number of mail groups for printing to separate Text files

6.19.3   Mail Printing to HTML file

By default Mbox Viewer supports printing of all mails, single or selected group of mails to single HTML file. Mails can also be printed to a separate per mail HTML file for further processing as described in "Mail Printing to PDF" 6.19.5 section and "Mail Printing to PDF files by Power Users" 6.19.6 section.

All archive mails can be printed to single HTML file but HTML viewer tools/Web Browsers may not be able to process large HTML files, become very slow or completely overloaded.

User can select and print groups of mails to separate HTML files as a work around. There is no option to automatically split all mails into configurable number of mail groups for printing to separate HTML files.

There is no hard rule how many mails can be printed to a single HTML file and viewed by a Web Browser. It depends on the size and content (such as heavy graphics) of mails. It should be doable to

print up to a couple of thousands small to medium size text mails to a single HTML file for further processing.

6.19.4　　Mail Printing to PDF printer

Mbox Viewer supports printing of all, single or selected group of mails to PDF printer.

Printing to PDF printer relies on Microsoft HTML Document object to print its content. First, Mbox Viewer will print mails to a single HTML file, then it will load this file into HTML Document object and request the document object to print itself.

The limitations as far as the maximum number of mails that can be printed are described in the "Printing Mails to HTML file" section.

The "File -> Print Config -> Page Setup" dialog option allows users to control the page title, header, footer and background color.

By default user is prompted to select PDF printer for printing mails. If the PDF printer is configured as the default printer, user can set "File -> Print Config -> Do Not Prompt" dialog option to skip the printer prompt step.

6.19.5　　Mail Printing to PDF file

Mbox Viewer supports multiple ways to print all, single or multiple mails to PDF file. All methods convert HTML files, created first from mails by the Mbox Viewer, to PDF format.

By default all, single or selected mails are printed to a single HTML file and then converted to PDF format. The default method limits the number of mails that can be effectively printed to single HTML file as noted in previous sections. More scalable method for is described in "Mail Printing to PDF files by Power Users" 6.19.6 section.

Mbox Viewer supports an option to print mails directly to PDF file without any user interaction. Direct printing to PDF file invokes external application to convert HTML files to PDF. By default the standard Google Chrome browser in so called headless mode is invoked to perform conversion. The second option is provided to leverage free wkhtmltopdf application to covert HTML to PDF.

Option to leverage the wkhtmltopdf is provided since the Chrome browser in the headless mode doesn't support options to control the page title, header and footer and the background color. It also enables users to run custom script with different HTML to PDF converter.

User can set "File -> Print Config -> Path To User Defined Script" to invoke HTML2PDF-single-wkhtmltopdf.cmd script included in the release package. HTML2PDF-single-wkhtmltopdf.cmd script creates PDF files with the right footer "Page Number of Total Pages" and no page Title and Header. The script can be customized by users if desired.

The wkhtmltopdf can be downloaded from <u>wkhtmltopdf.org downloads</u>.
The wkhtmltopdf command line options are documented in <u>wkhtmltopdf.org usage</u>.

User can replicate HTML2PDF-single-wkhtmltopdf.cmd to a new file and customize to leverage

different converter possibly commercial. User will need to update "Path To User Defined Script" in "File -> Print Config" dialog to point to new script path.

When printing multiple mails to a single PDF file, user can set "File -> Print Config -> Page Break After Each Mail" option to start each mail at new page.

User can unset "File -> Print Config -> Mail Hdr Background Color" option to remove mail header background color and save on the ink when printing. Useful option when using Chrome browser to print.

Lastly, HTML files can be open within Firefox, Chrome, IE, etc. Browser and printed to PDF file.

6.19.6   Mail Printing to PDF files by Power Users

Some users may need to print larger number of mails to a single PDF file. This is supported by Mbox Viewer but one or multiple manual steps are required. User must first print all mails to separate PDF files and then merge them into a single PDF file by running provided script.

In order to print mails to separate PDF files, user must first set "File -> Print Config -> Print mails to separate files -> PDF" option. After that, using standard mail print to PDF options, user can print all or large subset of mails to separate PDF files into sub-folder PDF_GROUP created by Mbox Viewer.

Printing large number of mails to separate PDF files is time consuming. Depending on the size and content of a single mail, it may take fraction of a second to create PDF file or tens of seconds.

PDFMerge-pdfbox.cmd script included in the release package can be used to merge PDF files. PDFMerge-pdfbox.cmd script invokes free PDFBox java tool to merge PDF files.

Free PDFBox java tool can be downloaded from pdfbox.apache.org.

PDFBox command line tools usage example can be found in pdfbox.apache.org commandline tools.

Java 8 can must be installed and can be downloaded from jdk8-downloads.

PDFMerge-pdfbox.cmd script must be copied into the directory housing all per mail PDF files and executed. PDFMerge-pdfbox.cmd script will create one or more merged PDF files in the PDF_MERGE sub-folder. If more than one merged PDF file is created in the PDF_MERGE sub-folder, PDFMerge-pdfbox.cmd script needs to be copied to that sub-folder and script executed again. The above steps need to be repeated until PDFMerge-pdfbox.cmd script creates single PDF file only.

Printing mails to separate PDF files can also be done offline. User must first set "File -> Print Config -> Print mails to separate files -> HTML" option. After that, using standard mail print to HTML options, user can print all or large subset of mails to separate HTML files into sub-folder HTML_GROUP created by Mbox Viewer.

HTML2PDF-all-chrome.cmd or HTML2PDF-all-wkhtmltopdf.cmd script included in the release package can then be used to convert HTML files to PDF files. One of these scripts must be copied into the directory housing all per mail HTML files, i.e. HTML_GROUP, and executed. Created PDF

files can then be merged into a single PDF file using PDFMerge-pdfbox.cmd script as described above.

6.19.7   Print Directory Structure

Mbox Viewer creates and manages several sub-directories in the directory that houses one or more mail archive files.

```
ROOT_MAIL_DIRECTORY
ROOT_MAIL_DIRECTORY\MailArchiveFile1.mbox
ROOT_MAIL_DIRECTORY\MailArchiveFile2.mbox
ROOT_MAIL_DIRECTORY\ImageCache
ROOT_MAIL_DIRECTORY\ImageCache\MailArchiveFile1- target directory for image files, such as png, jpg, etc, embeded into mails
ROOT_MAIL_DIRECTORY\ImageCache\MailArchiveFile2
ROOT_MAIL_DIRECTORY\ArchiveCache
ROOT_MAIL_DIRECTORY\ArchiveCache\MailArchiveFile1- target directory for saving Find Mails and User Selected Mails as .mbox archive files
and .mboxlist mail list files
ROOT_MAIL_DIRECTORY\Archive\MailArchiveFile2
ROOT_MAIL_DIRECTORY\PrintCache
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1 - target directory for printing to single CSV, TEXT, HTML and PDF files
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1\PDF_GROUP - target directory for printing to separate PDF files
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1\PDF_GROUP\PDF_MERGE - target directory for merged PDF files
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1\PDF_GROUP\PDF_MERGE\PDF_MERGE - target directory for merged PDF files if
multiple merge steps are needed
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1\HTML_GROUP - target directory for printing to separate HTML files
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1\HTML_GROUP\PDF_MERGE - target directory for merged PDF files
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile1\HTML_GROUP\PDF_MERGE\PDF_MERGE - target directory for merged PDF files if
multiple merge steps are needed
ROOT_MAIL_DIRECTORY\PrintCache\MailArchiveFile2
```

Mbox Viewer also creates and manages temporary directory C:\Users**UserName**\AppData\Local\ Temp\mboxview directory to store temporary files (attachments, eml and htm) created when a mail is selected by the user.

7 Windows Code Page Identifiers

The table below shows the supported character sets/code pages by Windows platform. The list is based on the following link.

https://msdn.microsoft.com/en-us/library/windows/desktop/dd317756(v=vs.85).aspx

Code Page	Name	Display Name
37	IBM037	IBM EBCDIC (US-Canada)
437	IBM437	OEM United States
500	IBM500	IBM EBCDIC (International)
708	ASMO-708	Arabic (ASMO 708)
720	DOS-720	Arabic (DOS)
737	ibm737	Greek (DOS)
775	ibm775	Baltic (DOS)
850	ibm850	Western European (DOS)

Code Page	Name	Display Name
852	ibm852	Central European (DOS)
855	IBM855	OEM Cyrillic
857	ibm857	Turkish (DOS)
858	IBM00858	OEM Multilingual Latin I
860	IBM860	Portuguese (DOS)
861	ibm861	Icelandic (DOS)
862	DOS-862	Hebrew (DOS)
863	IBM863	French Canadian (DOS)
864	IBM864	Arabic (864)
865	IBM865	Nordic (DOS)
866	cp866	Cyrillic (DOS)
869	ibm869	Greek , Modern (DOS)
870	IBM870	IBM EBCDIC (Multilingual Latin-2)
874	windows-874	Thai (Windows)
875	cp875	IBM EBCDIC (Greek Modern)
932	shift_jis	Japanese (Shift-JIS)
936	gb2312	Chinese Simplified (GB2312)
949	ks_c_5601-1987	Korean
950	big5	Chinese Traditional (Big5)
1026	IBM1026	IBM EBCDIC (Turkish Latin-5)
1047	IBM01047	IBM Latin-1
1140	IBM01140	IBM EBCDIC (US-Canada-Euro)
1141	IBM01141	IBM EBCDIC (Germany-Euro)
1142	IBM01142	IBM EBCDIC (Denmark-Norway-Euro)
1143	IBM01143	IBM EBCDIC (Finland-Sweden-Euro)
1144	IBM01144	IBM EBCDIC (Italy-Euro)
1145	IBM01145	IBM EBCDIC (Spain-Euro)
1146	IBM01146	IBM EBCDIC (UK-Euro)
1147	IBM01147	IBM EBCDIC (France-Euro)
1148	IBM01148	IBM EBCDIC (International-Euro)
1149	IBM01149	IBM EBCDIC (Icelandic-Euro)
1200	utf-16	Unicode
1201	unicodeFFFE	Unicode (Big-Endian)
1250	windows-1250	Central European (Windows)
1251	windows-1251	Cyrillic (Windows)
1252	Windows-1252	Western European (Windows)
1253	windows-1253	Greek (Windows)
1254	windows-1254	Turkish (Windows)
1255	windows-1255	Hebrew (Windows)

Code Page	Name	Display Name
1256	windows-1256	Arabic (Windows)
1257	windows-1257	Baltic (Windows)
1258	windows-1258	Vietnamese (Windows)
1361	Johab	Korean (Johab)
10000	macintosh	Western European (Mac)
10001	x-mac-japanese	Japanese (Mac)
10002	x-mac-chinesetrad	Chinese Traditional (Mac)
10003	x-mac-korean	Korean (Mac)
10004	x-mac-arabic	Arabic (Mac)
10005	x-mac-hebrew	Hebrew (Mac)
10006	x-mac-greek	Greek (Mac)
10007	x-mac-cyrillic	Cyrillic (Mac)
10008	x-mac-chinesesimp	Chinese Simplified (Mac)
10010	x-mac-romanian	Romanian (Mac)
10017	x-mac-ukrainian	Ukrainian (Mac)
10021	x-mac-thai	Thai (Mac)
10029	x-mac-ce	Central European (Mac)
10079	x-mac-icelandic	Icelandic (Mac)
10081	x-mac-turkish	Turkish (Mac)
10082	x-mac-croatian	Croatian (Mac)
12000	utf-32	Unicode (UTF-32)
12001	utf-32BE	Unicode (UTF-32 Big-Endian)
20000	x-Chinese-CNS	Chinese Traditional (CNS)
20001	x-cp20001	TCA Taiwan
20002	x-Chinese-Eten	Chinese Traditional (Eten)
20003	x-cp20003	IBM5550 Taiwan
20004	x-cp20004	TeleText Taiwan
20005	x-cp20005	Wang Taiwan
20105	x-IA5	Western European (IA5)
20106	x-IA5-German	German (IA5)
20107	x-IA5-Swedish	Swedish (IA5)
20108	x-IA5-Norwegian	Norwegian (IA5)
20127	us-ascii	US-ASCII
20261	x-cp20261	T.61
20269	x-cp20269	ISO-6937
20273	IBM273	IBM EBCDIC (Germany)
20277	IBM277	IBM EBCDIC (Denmark-Norway)
20278	IBM278	IBM EBCDIC (Finland-Sweden)
20280	IBM280	IBM EBCDIC (Italy)

Code Page	Name	Display Name
20284	IBM284	IBM EBCDIC (Spain)
20285	IBM285	IBM EBCDIC (UK)
20290	IBM290	IBM EBCDIC (Japanese katakana)
20297	IBM297	IBM EBCDIC (France)
20420	IBM420	IBM EBCDIC (Arabic)
20423	IBM423	IBM EBCDIC (Greek)
20424	IBM424	IBM EBCDIC (Hebrew)
20833	x-EBCDIC-KoreanExtended	IBM EBCDIC (Korean Extended)
20838	IBM-Thai	IBM EBCDIC (Thai)
20866	koi8-r	Cyrillic (KOI8-R)
20871	IBM871	IBM EBCDIC (Icelandic)
20880	IBM880	IBM EBCDIC (Cyrillic Russian)
20905	IBM905	IBM EBCDIC (Turkish)
20924	IBM00924	IBM Latin-1
20932	EUC-JP	Japanese (JIS 0208-1990 and 0212-1990)
20936	x-cp20936	Chinese Simplified (GB2312-80)
20949	x-cp20949	Korean Wansung
21025	cp1025	IBM EBCDIC (Cyrillic Serbian-Bulgarian)
21866	koi8-u	Cyrillic (KOI8-U)
28591	iso-8859-1	Western European (ISO)
28592	iso-8859-2	Central European (ISO)
28593	iso-8859-3	Latin 3 (ISO)
28594	iso-8859-4	Baltic (ISO)
28595	iso-8859-5	Cyrillic (ISO)
28596	iso-8859-6	Arabic (ISO)
28597	iso-8859-7	Greek (ISO)
28598	iso-8859-8	Hebrew (ISO-Visual)
28599	iso-8859-9	Turkish (ISO)
28603	iso-8859-13	Estonian (ISO)
28605	iso-8859-15	Latin 9 (ISO)
29001	x-Europa	Europa
38598	iso-8859-8-i	Hebrew (ISO-Logical)
50220	iso-2022-jp	Japanese (JIS)
50221	csISO2022JP	Japanese (JIS-Allow 1 byte Kana)
50222	iso-2022-jp	Japanese (JIS-Allow 1 byte Kana - SO/SI)
50225	iso-2022-kr	Korean (ISO)
50227	x-cp50227	Chinese Simplified (ISO-2022)
51932	euc-jp	Japanese (EUC)
51936	EUC-CN	Chinese Simplified (EUC)

Code Page	Name	Display Name
51949	euc-kr	Korean (EUC)
52936	hz-gb-2312	Chinese Simplified (HZ)
54936	GB18030	Chinese Simplified (GB18030)
57002	x-iscii-de	ISCII Devanagari
57003	x-iscii-be	ISCII Bengali
57004	x-iscii-ta	ISCII Tamil
57005	x-iscii-te	ISCII Telugu
57006	x-iscii-as	ISCII Assamese
57007	x-iscii-or	ISCII Oriya
57008	x-iscii-ka	ISCII Kannada
57009	x-iscii-ma	ISCII Malayalam
57010	x-iscii-gu	ISCII Gujarati
57011	x-iscii-pa	ISCII Punjabi
65000	utf-7	Unicode (UTF-7)
65001	utf-8	Unicode (UTF-8)

7.1 Handling Unknown Code Page Names

In a few cases, mboxview will map unknown code names to the equivalent code names supported by Windows.

Unknown Code Name	Mapped To Code Name
`gbk`	`GB18030`
`cp1252`	`Windows-1252`
`ascii`	`us-ascii`
`cp819`	`iso-8859-1`
`latin1`	`iso-8859-1`
`latin2`	`iso-8859-2`
`big-5`	`big5`
`SJIS`	`shift_jis`