# 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 text, attachment names and attachment text 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.
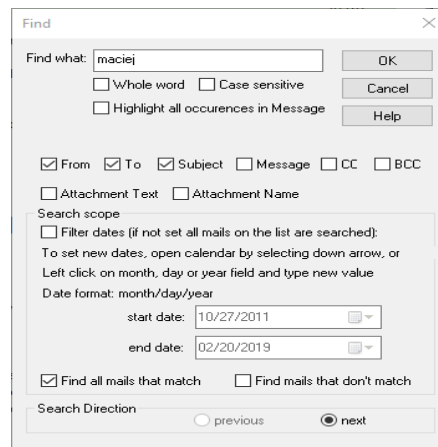


To repeat the search in order to find the next matching mail,  click the "Edit" drop menu and select "Find again" option, or select F3 key.

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

## 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.



## 1.2  Find mails that don't match option

User can check "Find mails that don't match" option to find mails that don't match search criteria.

## 1.3  Handling of special '*' search string

User can specify single '*' character as the search string to find subset of mails as follows:

1. Find mails that have CC header field by checking out CC check box only.

2. Find mails that have BCC header field by checking only BCC check box only.

3. Find mails that have at least one attachment by checking out Attachment Name check box only.

4. Match all mails by checking out any of others check boxes only.

NOTE that if the Find all occurrences option is not selected and Filter Dates option is no set, user will be traversing the subset of mails between the start and end dates.

## 1.4  Filter Dates

Filter Dates option enables users to search mails specified by the date range. By default the start date corresponds to the oldest mail and the end date corresponds to the latest mail. If the Filter Dates option is not set, all mails on the selected list (All Mails, Find Mails, User Selected Mails) are searched regardless of start and end dates values. If the Filter Dates option is set, user can configure the start and

end dates for searching. The start and end date processing depends on the mail list selected (All Mails, Find Mails, User Selected Mails). While in All Mails list, the start and end dates are fully under the user control regardless of other user actions such as selecting other lists. However, the start and dates do not persist across application restarts or when new mbox mail archive file is selected by the user. User can set the start and end dates while in Find Mails or User Selected Mails list but the initial values are set to default values.
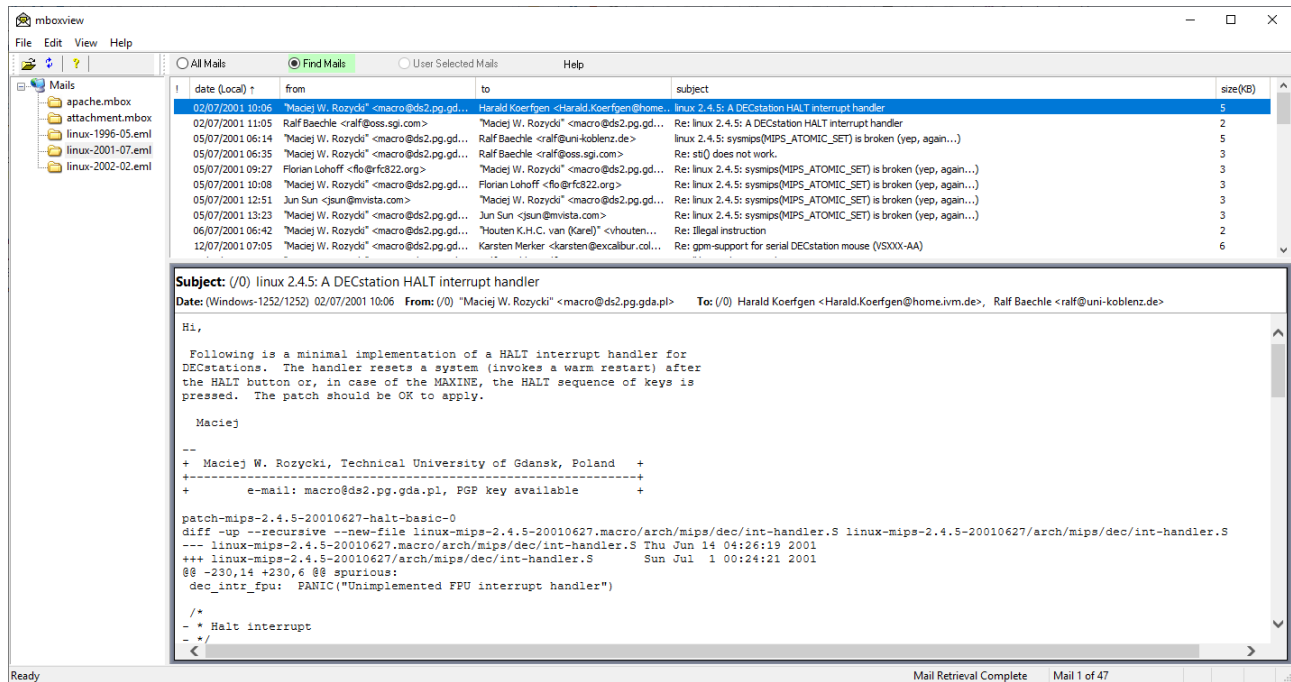


## 2   Advanced Find

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

Users can configure the search scope by selecting all or subset of the from, to, subject, cc, bcc, message text, attachment name 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 cases. Select "Configure Filter Rule" button to start dialog and select one of the preconfigured filter rules.



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.

## 2.1  Find all mails that don't match option

User can check "Find all mails that don't match" option to find mails that don't match search criteria.

## 2.2  Handling of special '*' search string

User can specify single '*' character as the search string in any of the Filter fields to find subset of mails as follows:

1. Find all mails that have CC header field by checking out CC check box only.

2. Find all mails that have BCC header field by checking only BCC check box only.

3. Find all mails that have at least one attachment by checking out Attachment Name check box only.

4. Match all mails by checking out any of others check boxes only.

## 2.3  Filter Dates

Filter Dates processing is the same for Basic Find and Find Advanced search. See chapter 1.4.

# 3    Search Reliability

MBox Viewer searching is Unicode based since version 1.0.3.40.  The search string, mail header fields and mail content are encoded first as Unicode UTF8 strings. Encoding into UT8 is done on the fly in most cases.

Comparing Unicode strings has challenges since the same character may have multiple binary representation. To overcome the problem, Unicode strings need to be normalized before they can be compared.   MBox Viewer relies on non-canonical normalization form C. String normalization is not without issues. **To help to resolve or minimize issues, please create tickets and provide as many details as possible to help to diagnose the issue.**

## 3.1  Search for Words

Search for words is quite reliable for words composed from the following characters from ASCII and ISO-8859-1 (Also known as ISO Latin 1) character sets:

```
"0123456789abcdefghijklmnopqrstuvwxyz_ABCDEFGHIJKLMNOPQRSTUVWXYZÀÁÂÃÄÅÆÇÈÉÊËÌÍÎÏÐÑÒÓÔÕÖØÙ
ÚÛÜÝÞßàáâãäåæçèéêëìíîïðñòóôõöøùúûüýþÿ"
```

Any other characters from ASCII and ISO-8859-1 sets are considered as word delimiters.

Searching for words in other character sets/languages is currently not reliable for the following reasons:

1. Some languages don't have word delimiters

2. Set of all UNICODE delimiters is very large and currently not known to MBox Viewer. Future releases may enhance the search.

Option to search for words should not be set when searching emails not composed from ASCII and ISO-8859-1 character sets.