

EMC Documentum 6.5 SP2 Email Search Applications

Applied Technology

Abstract

This white paper describes how to customize EMC® Documentum® 6.5 SP2 applications to search email attributes such as from, to, cc, and bcc, email distribution lists, and attachment formats and names. Samples are provided for DFC, Webtop, and DFS services.

May 2009

Copyright © 2009 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

Part Number h6339

Table of Contents

Executive summary	4
Introduction	4
Audience	4
Terminology	4
Email support	4
Alias mapping	5
Webtop configuration	6
Enabling dm_message_archive	6
Search on additional email attributes	7
Configure email attributes to be displayed	9
Troubleshooting	11
DFC email queries	11
DFS email queries	12
Conclusion	13
References	13

Executive summary

EMC® Documentum® search applications are based on searching for an object type based on its full-text content and/or attributes. Emails are represented by a complex object type, `dm_message_archive`. Webtop allows users to search on the `dm_message_archive` attributes such as subject, to, from, and sent and received dates. This white paper shows you how to configure email searches on any email attributes.

Documentum provides a mapping file with aliases that support searching the following additional email attributes and content: email body, CC field, BCC field, DIST field (distribution list), any recipient, attachment format, attachment name, and attachment content. These aliases can be used with DFC, DFS, and Webtop queries.

Introduction

This white paper is organized in the following order:

- Email support
- Alias mapping
- Webtop configuration
- Documentum Foundation Classes (DFC) email queries
- Documentum Foundation Services (DFS) email queries

This paper briefly describes the Documentum features enabled by the `dm_message_archive` email format. For Webtop, an example is provided for configuring the application display additional email attributes in search results. Examples of customized search in Webtop, DFC, and DFS use aliases to query and display complex object attributes.

Audience

This white paper assumes that for custom DFC or DFS applications you are familiar with the Documentum object model, Java programming language, and DFC, DFS, or Webtop customization models. For Webtop configuration, Java is not a requirement.

Terminology

DFC: DFC provides a framework of Java classes that gives access to Content Server functionality through custom DFC, Webtop, and DFS applications. The DFC query builder API is used to build complex queries in an application. The example in this white paper uses the DFC query builder API.

DFS: DFS is a set of technologies that enables service-oriented, programmatic access to the EMC Documentum Content Server platform and related products. The search service is an Enterprise Content Service built on DFS. The example in this white paper builds a structured query.

Webtop: Webtop is an EMC Documentum web application that provides access to Content Server functionality. Through the Webtop configuration framework, you can change functionality to suit your business needs. The example in this white paper uses XML and JSP standards to create custom email search components.

Email support

Webtop 6.5 SP1, DCO, and Documentum 5.3, 6.0, and 6.5 repositories support `dm_message_archive` objects in the proprietary EMCMF format. This format encapsulates the content, metadata, and attachment information as a single unit of information.

Content Server 6.5 supports the EMCMF mail message format. On versions 5.3 and 6.0 of Content Server, you must install supporting DocApps. Webtop 6.5 SP1 supports import in the new format. You must enable message archive support in app.xml and enable the installation of an email conversion application on each client. Refer to the *EMC Documentum Webtop Version 6.5 SP1 Email Migration Guide*, available on EMC Powerlink®, for instructions on installing support in repositories and configuring support in Webtop.

The following features have been enhanced to support the EMCMF format:

- **Import:** When a user imports an email message (.msg) in Webtop or DCO, the email message can be converted to EMCMF format and stored as a dm_message_archive type. All attachments in the email are imported and related to the email.
- **Export:** When a user exports an EMCMF object in Webtop or DCO, the object can be converted to native email format (.msg) for viewing by Microsoft Outlook.
- **View Properties and Listing pages:** The Properties and Listing pages in Webtop have been enhanced to show email-centric attributes such as To, From, Date Sent, and Subject.
- **Transform:** EMCMF messages can be transformed to HTML, XML, or PDF.

Existing email messages of the dm_email_message object type (and its subtypes) and the msg format must be migrated to the dm_message_archive type (EMCMF) or one of its subtypes. EMC Documentum provides a utility to perform the migration of dm_email_message objects to dm_message_archive objects. This utility can be executed as many times as required. They report if there are any dm_email_message objects that need to be migrated. The *EMC Documentum Webtop Version 6.5 SP1 Email Migration Guide* has further details on these utilities.

Alias mapping

Complex object attributes are represented by a path built using a mapping definition file. For example, the path for the sender attribute is messageData/**fromAddress**/friendly_name. The path for the recipients attribute is messageData/**toAddress**/friendly_name. These aliases can be used in Webtop configuration and in DFC and DFS custom applications.

Table 1 lists available aliases. These aliases will be used in the examples of this white paper.

Table 1. Available aliases

Alias	Description
messageData/fromAddress/email_addr messageData/fromAddress/friendly_name	Property alias: From field
messageData/toAddress/email_addr messageData/toAddress/friendly_name	Property alias: To field
messageData/ccAddress/email_addr messageData/ccAddress/friendly_name	Property alias: cc field
messageData/bccAddress/email_addr messageData/bccAddress/friendly_name	Property alias: bcc field
messageData/distAddress/email_addr messageData/distAddress/friendly_name	Property alias: dist field
messageData/recipientAddress/email_addr messageData/recipientAddress/friendly_name	Property alias: Any recipient (to, cc, ...)
messageData/allAddresses/email_addr messageData/allAddresses/friendly_name	Property alias: Any address (from, to, cc, ...)
messageAttachment/format_name	Property alias: Format name of the attachment. Use Documentum format name (ex.: msg)
messageAttachment/attachmentName	Property alias: Name of the attachment
messageAttachment/attachmentContent/ft	Content alias: Full-text content of the attachment
body	Content alias: Full-text content of the email body
embedded_body	Content alias: Full-text content of an email embedded in the main email

Note: Some of these aliases are new in 6.5 SP2, so they will not work with previous versions of Webtop. The last three aliases can only be used in full-text queries. They cannot be used for database-only searches, for example, if the DFC query builder flag `setDatabaseSearch` is set to true.

Webtop configuration

WDK-based applications of version 6.5 SP1 or later support viewing and exporting `dm_message_archive` objects. Users can view or export Outlook mail messages as .msg files in HTML. To view them as `dm_message_archive` objects with their complex attributes, the client must have a converter to convert `dm_message_archive` objects back to .msg files. This converter can be automatically installed as part of the UCF download. However, the converter can take a long time to download and install. You can avoid this delay by deploying the `ES1_MRE.exe` installer using standard mechanisms, such as Microsoft Systems Management Server (SMS).

Enabling dm_message_archive

To support the `dm_message_archive` type, you must enable it in `app.xml` and install a message converter on the client. Your application must also be configured to use UCF, not HTTP, content transfer.

To enable a message, add the following lines to `app.xml` in your custom/config directory:

EMC Documentum 6.5 SP2
Email Search Applications
Applied Technology

```

<messageArchive-support>
  <enabled>true</enabled>
  <supported-file-format>
    <file-format>msg</file-format>
  </supported-file-format>
  <default-attachment-object-type>dm_document</default-attachment-object-type>
  <store-emf-object-as-archive>>false</store-emf-object-as-archive>
  <skip-duplicate-messages>true</skip-duplicate-messages>
</messageArchive-support>

```

Other configuration options are available in this XML node. For more information, refer to the *EMC Documentum Web Development Kit Development Guide*.

To enable download of the converter, uncomment the following lines in `wdk/contentXfer/ucf.installer.config.xml`:

```

<executables>
  <executable version="6.5.0"
    href="ES1_MRE.exe"
    installOptions='/qn ADDLOCAL=ALL'
    uninstallOptions='/x /qn REBOOT=REALLYSUPPRESS' />
</executables>

```

If your users have previously used Webtop for UCF content transfer, you will have to version `ucf.installer.config.xml` to force a download of the executable. Add a letter to the app element version value, for example:

```

<app id="shared" version="6.5.0.220a" compatibilityVersion="5.3.0" />

```

After you have successfully configured application support and restarted the application server, the EMCMF format should appear as an option in the import dialog, as shown in Figure 1.

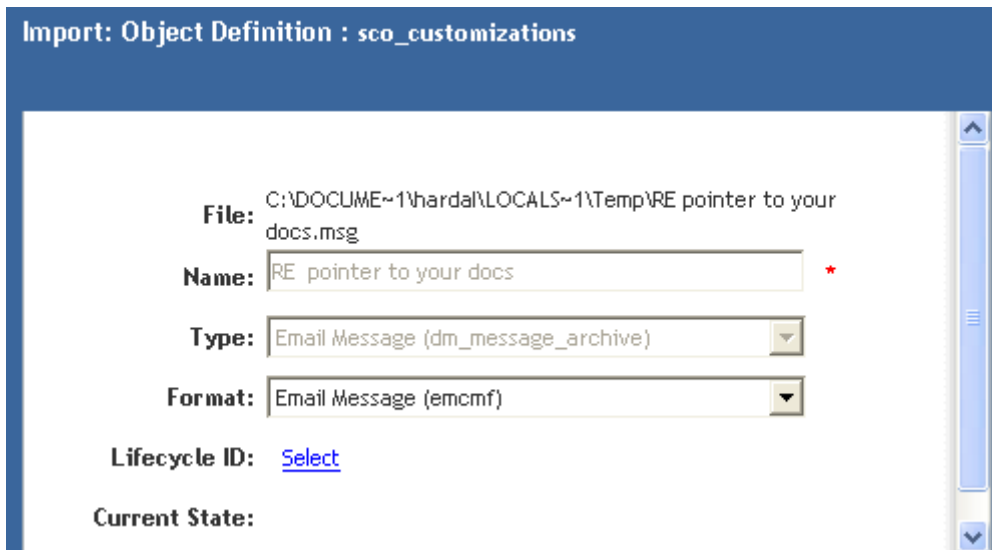


Figure 1. The EMCMF format is now available

Search on additional email attributes

If email messages in your repository have been migrated to the `dm_message_archive` type, Webtop users can select this type and search for some email attributes out of the box: Subject, To, From, and Sent and Received dates. When the user selects the `dm_message_archive` type, the following search screen is displayed as shown in Figure 2:

Figure 2. Search screen

You can add or remove email attributes in the email search screen by substituting your own JSP page in place of the one shown above. To do this, perform the following steps.

1. Modify the XML configuration file for the advanced search component of type `dm_message_archive`. Create a file named `messagearchive_advsearchex_modifications.xml` in the directory `custom/config` with the following content:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<config version='1.0'>
<scope type="dm_message_archive">
<component modifies="advsearch:webtop/config/messagearchive_advsearchex_component.xml">
  <replace path="pages.start">
    <start>/custom/messagearchive/messagearchive_advsearchex.jsp
    </start>
  </replace>
</component>
</scope>
</config>
```

2. Copy the JSP page `/webcomponent/library/messagearchive/messagearchive_advsearchex.jsp` to `custom/messagearchive`. (Create the `messagearchive` directory if you have not already done so.)

- After the last table row (<tr>) before the closing dmfxs:repositorysearch tag, the row contains a dmfxs:searchdateattribute tag for receive_date — add the following row:

```
<tr>
<td class="leftAlignment" valign=top nowrap>
<dmfxs:searchattribute name='scoattrname' attribute="messageAttachment/attachment_name"
casevisible='false' width="<%=SearchAttribute.DEFAULT_FIELD_WIDTH%" />
</td>
</tr>
```

- Save the files and refresh memory on the application server by navigating to /wdk/refresh.jsp. When you go to the advanced search screen and select the dm_message_archive type, you see the new attribute, Attachment, displayed at the bottom:

Received: Anytime From Date To Date

Attachment name: contains indexing

Figure 3. Attachment attribute displays

If the user has selected the attachment name column to be displayed, a successful search will display the name as shown in the following search results for an attachment name containing “index”. (Configuration for email attributes such as attachment name is described in the next section.)

Name	CC	Ranking	Attachment	From	Sent
RE: config mbean	Fan, Zhenqiang	12%	ESSAdminIndex	Chen, Charlie	3/24/2009 8:25 PM

Figure 4. Attachment name column displays

Configure email attributes to be displayed

You can configure Webtop to display other attributes of dm_message_archive. The basic search component of type dm_message_archive in webtop/config/ messagearchive_search60_component.xml configures the columns to be displayed for email search results. Out of the box, the column display is similar to Figure 5.

Search Results
 9 Results in dm_notes - 3/31/09 3:11 PM

All Results (15) > 2009 Q1 (9)

Name	Ranking	From	Sent	To	Received	Importance
Message with attachments	54%	Administrator	12/31/79 4:00		11/8/07 5:25 AM	Normal
RE DFC 7 Check-In 55382 Done....	54%		nulldate		nulldate	
RE DCTM integration design spec...	54%		nulldate		nulldate	
RE Did anybody meet this proble...	53%		nulldate		nulldate	
RE CPS functional spec is DCTM-...	53%		nulldate		nulldate	
GorillaDeck-ISSCore-Text-Matt pp...	53%		nulldate		nulldate	
Prelim M4 results for single nod...	52%		nulldate		nulldate	
Checking to see if the Developer...	51%		nulldate		nulldate	
[Webtop-review-3] All Reviewers ...	50%	Anurag	2/3/09 4:29	Totale, Sachin	2/3/09 4:29 AM	Normal

Figure 5. Column display

Two of these columns—**From** and **To**—use aliases to map the complex object attributes. You can display any attributes by using aliases from the alias mapping file.

The following example adds the CC field and the name of the attachment. To do this, perform the following steps:

1. Modify the XML configuration file for the search component of type dm_message_archive by creating a file named messagearchive_search60_modifications.xml in the directory custom/config with the following content:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<config version='1.0'>
<scope type="dm_message_archive">
<component modifies="search:webtop/config/messagearchive_search60_component.xml">
  <insert path="columns_list">
    <column>
      <attribute>messageData/ccAddress/friendly_name
    </attribute>
    <label>CC</label>
    <visible>>true</visible>
    </column>
    <column>
      <attribute>messageAttachment/attachmentName
    </attribute>
    <label>Attachment</label>
    <visible>>true</visible>
    </column>
  </insert>
</component>
</scope>
</config>
```

2. Refresh the components in the app server memory by navigating to wdk/refresh.jsp. Then you can select these columns for display in the search results. (If you have already performed a search, you must launch it again for the new attribute values to be displayed.) The additional attributes are displayed as shown in Figure 6.

Search Results
 25 Results in winsql_213 - 4/28/2009 4:45 PM

Name	Attachment	From	Sent	To	CC
RE: Bug 168050		Rajadurai, Thangam	10/29/2008 12:37 AM	Muslim, Aejaz,Sateesh, Arvapally	Totale, Sachin
Appreciation email from John		Durairaj, Visva	3/11/2009 1:25 AM	India COE CMA KW Client Engg	
Does any one has hotfix for #		Zhang, Roy	3/10/2009 11:31 AM	webtop_wdk_support (DocumentL	
EMCMF_Differentformats_10	firefox_config_	Thilagam Pandian , HCI	3/4/2009 2:38 AM	Thilagam Pandian , HCL Software F	
testimport_with import	excel.xls		12/31/1979 4:00 PM	Kirubakaran GuthaM (HCL Finance	
testimport_without attachm			12/31/1979 4:00 PM	Thilagam Pandian , HCL Software F	
RE: more than one analyzer p		Natarajan, Valli	4/21/2009 12:50 AM	Hardage-Vergeer, Lani,Chandrase	
Yet Another Round of Inform	image001.png	Zhong, Jove	4/10/2009 3:31 AM	Chen, Noah,Snow, Evan,Fan, Zhen	Farooq, Aamir

Figure 6. Additional attributes displayed

Troubleshooting

Check the following:

- In the custom app.xml, is message_archive enabled?
- Is the message converter download enabled in ucf.installer.config.xml?
- Does the client host have DLLS files in C:\Program Files\Common Files\EMC, indicating a successful download of the converter?
- Can you find an email if you search for it?
- Did you choose the EMCMF format when you imported the message?
- Did you wait long enough for the email message to be indexed and available for search?
- Try to get the message using IAPI or IDQL. Substitute a word from the message subject in place of <message subject> in the following example:

```
select r_object_id, object_name from dm_message_archive where
message_subject like '%<message subject>%'
```

DFC email queries

In the DFC query model, you construct a query using the APIs of IDfQueryBuilder. IDfQueryBuilder is in the package com.documentum.fc.client.search.

In the addResultAttribute() and addSimpleAttrExpression() methods, you can add aliases to email attributes that are not attributes of dm_message_archive. Use the aliases in IDfResultEntry.get... methods to get aliased attribute results.

In the following example, the To address is added to the query and to the results in the highlighted lines. The To address is retrieved in the results at the end of this code snippet:

```
IDfClient client = DfClient.getLocalClient();
m_sessionManager = client.newSessionManager();
DfLoginInfo identity = new DfLoginInfo(USER, PASSWORD);
m_sessionManager.setIdentity(DOCBASE, identity);

m_searchService = client.newSearchService(m_sessionManager, DOCBASE);
IDfQueryManager queryManager = m_searchService.newQueryMgr();
m_queryBuilder = queryManager.newQueryBuilder("dm_message_archive");
m_queryBuilder.addSelectedSource(DOCBASE);
```

```

m_queryBuilder.addResultAttribute("object_name");
m_queryBuilder.addResultAttribute("message_subject");
m_queryBuilder.addResultAttribute("receive_date");
m_queryBuilder.addResultAttribute("message_size");
m_queryBuilder.addResultAttribute("messageData/toAddress/email_addr");

IDfExpressionSet exprSet = m_queryBuilder.getRootExpressionSet();
exprSet.addFullTextExpression("test");
exprSet.addSimpleAttrExpression("messageData/toAddress/email_addr",
    IDfAttr.DM_STRING,
    IDfSimpleAttrExpression.SEARCH_OP_CONTAINS, false, false, "neal");

IDfQueryProcessor processor = m_searchService.newQueryProcessor(m_queryBuilder,
true);
processor.blockingSearch(SEARCH_TIMEOUT);
IDfResultsSet results = processor.getResults();
for (int i = 0 ; i < results.size(); i++)
{
    IDfResultEntry result = results.getResultAt(i);
    System.out.println("title = " + result.getString("object_name"));
    System.out.println("recipient[0] = " +
        result.getRepeatingString("messageData/toAddress/email_addr", 0));
}

```

For more information on constructing queries in DFC, refer to the DFC chapter in *EMC Documentum Search Development Guide*, available on Powerlink.

DFS email queries

The DFS search service provides full-text and structured search against multiple EMC Documentum repositories and external sources. For details on the search service, refer to the *EMC Documentum Enterprise Content Services Reference*, available on Powerlink.

A structured query object defines a query using an object-oriented model. The query is constrained by a set of criteria in an ExpressionSet object, and the scope of the query (sources against which the query runs) is defined by an ordered list of RepositoryScope objects.

The query criteria are defined in an ExpressionSet object, which consists of Expression objects. Each Expression object defines either a full-text expression or a search constraint on a single object property. Expression instances are related by a single logical operator, AND or OR. An ExpressionSet can nest ExpressionSet instances to form complex expression trees.

Three classes extend the Expression class: FullTextExpression, PropertyExpression, and ExpressionSet. All three will be used in the example code. Aliases for dm_message_archive attributes can be used in building an expression set. FullTextExpression can take content aliases, and PropertyExpression can take property aliases.

The following example from a DFS customization creates a structured query instance with the dm_message_archive object type, sets the repositories, and adds expressions to the expression set. The properties to be displayed are defined in a PropertyFile object.

```

StructuredQuery query = new StructuredQuery();
query.setObjectType("dm_message_archive");
query.setRepositories(Arrays.asList(DOCBASE));

ExpressionSet expression = new ExpressionSet(ExpressionSetOperator.AND);
expression.addExpression(new FullTextExpression("test"));
expression.addExpression(new

```

```
PropertyExpression("messageData/toAddress/email_addr", Condition.CONTAINS,
"neal"));
query.setRootExpressionSet(expression);

ServiceFactory serviceFactory = ServiceFactory.getInstance();
ISearchService searchService = serviceFactory.getService(ISearchService.class,
ContextFactory.getInstance().getContext());

PropertyProfile propProfile = new PropertyProfile();
propProfile.setFilterMode(PropertyFilterMode.SPECIFIED_BY_INCLUDE);
propProfile.setIncludeProperties(Arrays.asList("object_name",
"messageData/toAddress/email_addr"));

OperationOptions options = new OperationOptions();
options.setPropertyProfile(propProfile);

QueryExecution queryExecution = new QueryExecution(0, 100, 100);
// Execute search
QueryResult queryResult = searchService.execute(query, queryExecution, options);
//...
```

Conclusion

Custom email search components can facilitate the search process for users and enhance return on investment. Documentum search customizations based on WDK, DFC, or DFS can create queries and display the properties of the complex object `dm_message_archive`.

References

For more information on DFC, DFS and WDK/Webtop programming, refer to the following development guides on Powerlink:

- *EMC Documentum Foundation Classes Development Guide*
- *EMC Documentum Foundation Classes javadocs*
- *EMC Documentum Foundation Services Development Guide*
- *EMC Documentum Web Development Kit Development Guide*