

IBM Content Collector
Version 4.0.1.13

Db2 Text Search guide

© International Business Machines Corporation, 2021. All rights reserved.

What's new in Content Collector 4.0.1.13

The IBM Content Collector fix pack 13 installation supports the following basic scenarios:

1. Fresh ICC installation with Db2TS

Prerequisite – Install Db2 as custom installation selecting Db2 Text Search feature. Install and configure the CM8 server choosing Enable Db2 Text Search option during profile configuration.

With IBM Content Collector FP13, customer can use Db2 Text Search indexing to index the documents archived in Content Manager repository. Install Db2 Text Search Support Installer of ICC on the CM8 server enabled for Text Search. After completing the configuration with Db2TS, the customer needs to create a new item type to this CM8 server for archival and indexing of the documents.

2. Install FP13 and transition to Db2TS

Prerequisite – Re-install Db2 as custom installation selecting Db2 Text Search feature. Re-configure the existing CM8 server profile of NSE, selecting Enable Db2 Text Search option.

With IBM Content Collector FP13, to start indexing the documents archived in Content Manager repository, install Db2 Text Search Support Installer of ICC on the CM8 server enabled for Text Search. After completing the configuration with Db2TS, the customer needs to create a new item type to this CM8 server for archival and indexing of the documents and can continue searching against existing NSE and new Db2TS indexes.

3. Install FP13 and stay on NSE

Prerequisite – Existing CM8 server enabled with NSE indexing

In case the customer does not wish to use Db2TS indexing features provided with this fix pack, and only requires the fix pack installation for the security and APAR fixes then they can skip the Db2 Text Search Support installation/Configuration and continue using the CM8 Text Search Support installer of ICC for indexing with NSE. The customer can continue archiving, indexing, and searching against the existing item types. However, the creation of new item types is supported only with Db2TS indexing.

4. Install FP13, stay on NSE for a while and transition to Db2TS at a later point in time

Prerequisite – Existing CM8 server enabled with NSE indexing.

To switch to Db2TS later,

Prerequisite – Re-install Db2 as custom installation selecting Db2 Text Search feature. Re-configure the existing CM8 server profile of NSE, selecting Enable Db2 Text Search option.

In case the customer upgraded to Fix Pack 13 for the security and APAR fixes and continued using NSE indexing and now wishes to start indexing with Db2TS indexes, this can be achieved with Db2 Text Search Support installation on the Db2TS indexing enabled CM8 server. After completing the configuration with Db2TS, the customer needs to create a new item type to the CM8 server configured with Text Search.

Db2 Text Search Services environment

Ensure the CM8 environment has been set up with Text Search Indexing enabled.

Ensure Db2TS accessories suite has been installed and configured.

Installing Db2 Text Search Services Support in GUI mode

To install Db2 Text Search Services Support by using the installation wizard:

1. Stop the IBM Db2 Text Search Services instance in the CM8 server.
2. Download and extract the appropriate installation package of IBM Db2 Text Search service support where IBM Db2 Text Search Services is installed.
3. Provide the required environment variables:
For windows:
Default location of Db2HOME points to sqllib installation directory
Ex → Db2HOME = C:\Program Files\IBM\SQLLIB
For non-windows:
Default location of Db2HOME points to sqllib installation directory
Ex → export Db2HOME=/home/db2inst1/sqllib
Default location of IBMCMROOT points to db2cmv8 installation directory
Ex → export IBMCMROOT=/opt/IBM/db2cmv8/
4. Start the installer:
On windows, run install.exe.
On non-windows, run install.bin.
5. Answer the remaining prompts as the installer guides you through the installation process.
 - a. Specify an installation folder where Db2 Text Search Services Support is installed or use the default folder. The default path is C:\Program Files (x86)\ibm\ContentCollector_Db2TS_Support for Windows. The default path for non-windows is /opt/IBM/ContentCollector_Db2TS_Support.
 - b. Specify the root directory of the server where IBM Db2 Text Search Services config is installed. The default path is C:\ProgramData\IBM\Db2\Db2COPY1\Db2\db2tss for Windows. The default path for non-windows is /home/db2inst1/sqllib/db2tss/.
 - c. Specify the root directory of the server where IBM Db2 Text Search Services plugins is installed. The default path is C:\Program Files\IBM\SQLLIB\db2tss for Windows. The default path for non-windows is /opt/ibm/db2/V11.1/db2tss.
 - d. Select the source of the documents that will be marked for indexing.
Db2 Text Search Services Support modifies the constructor.xml file to deploy ICC constructors required depending on the source chosen.

Note: If the customer plans to update the constructor.xml manually, and does not require a connection to ICC, then during Db2TS Support installation, enter a dummy value for the host name of the ICC server running the IBM Content Collector web applications. In such cases, the constructor.xml file, needs to be updated manually post the completion of Db2TS support installation.

Important: For non-windows machines, provide full permissions recursively to the config/plugins directories so that the Db2TS Support installer can create/update the necessary files post installation.

Post-installation steps

Once the Db2 Text Search Services Support gets successfully installed, navigate to the installation directory < Db2TextSearchServicesSupportInstallDir>

For non-windows machines, the script needs to be run using the db2 instance user (ex- su db2inst1)

The installer deploys 2 scripts to insert the custom constructors of ICC for Email and SharePoint

1. If the source documents that you select when you install Db2 Text Search Service Support is Email, then run the AddICCConstructor_Email.ps1 on windows to insert the Email specific constructor to the database. For non-windows, the script deployed will be AddICCConstructor_Email.sh.
2. If the source documents that you select when you install Db2 Text Search Service Support is SharePoint, then run the AddICCConstructor_SharePoint.ps1 on windows to insert the SharePoint specific constructor to the database. For non-windows, the script deployed will be AddICCConstructor_SharePoint.sh.

In case the customer requires to index both Email and SharePoint, both the scripts must be run one after the other. These script when executed would request for the below details to establish a connection with the library server database

- Library Server database: icmnlbdb (l/p requested from the user)
- Library Server Username: icmadmin (l/p requested from the user)
- Library Server password (l/p requested from the user)

Please ensure that the Db2TS service is restarted as the script involves DB changes.

Important:

In case the script fails to execute successfully, the constructors can be inserted manually to the Content Manager repository through the Db2 Command Line or using SQL client.

Steps to manually insert the constructor to the database:

For non-windows machines, the queries need to be run using the db2 instance user (ex- su db2inst1)

EMAIL:

1. When using db2 command line, establish a connection with CM8
Connect to <library server database> user <library server user>
Hit enter and provide the password
2. Run the below query to verify if the constructor entry of the email already exists

```
SELECT COUNT(*) AS COUNT from SYSIBMTS.SYSTSDEFAULTS WHERE NAME = 'DATAPREP:AFUCTOR4EMAIL'
```
3. Run the below query to add the Email constructor entry if not already present.

```
INSERT INTO SYSIBMTS.SYSTSDEFAULTS VALUES('DATAPREP:AFUCTOR4EMAIL ', ", 1)
```
4. Run the below query to verify if the constructor has been inserted.

```
SELECT COUNT(*) AS COUNT from SYSIBMTS.SYSTSDEFAULTS WHERE NAME = 'DATAPREP:AFUCTOR4EMAIL'
```

SHAREPOINT:

1. When using db2 command line, establish a connection with CM8
Connect to <library server database> user <library server user>
Hit enter and provide the password

2. Run the below query to verify if the constructor entry of the SharePoint already exists

```
SELECT COUNT(*) AS COUNT from SYSIBMTS.SYSTSDEFAULTS WHERE NAME =  
'DATAPREP:AFUCTOR4SHAREPOINT'
```

3. Run the below query to add the SharePoint constructor entry if not already present.

```
INSERT INTO SYSIBMTS.SYSTSDEFAULTS VALUES('DATAPREP:AFUCTOR4SHAREPOINT'  
,", 1)
```

4. Run the below query to verify if the constructor has been inserted.

```
SELECT COUNT(*) AS COUNT from SYSIBMTS.SYSTSDEFAULTS WHERE NAME =  
'DATAPREP:AFUCTOR4SHAREPOINT'
```

Please note, For File Systems documents, IBM Content Collector uses the default constructor provided by Content Manager. The database entry for this constructor is added during Content Manager profile configuration and does not require explicit entry to be inserted.

Creating a CM8 item type for the first time in CM8 server

Additional configurational steps for item type creation

To be able to get the Db2TS service started successfully and create an item type successfully inserted into the repository, one of the below approaches needs to be followed:

- A. In case the CM8 server is freshly installed with Db2TS enabled with no prior NSE active catalog in ICC
 1. Update the constructor.xml manually OR
 2. Create a manual catalog and then proceed towards item type creation.
- B. 2. In case the CM8 server is upgraded from NSE to Db2TS with an existing NSE active catalog in ICC.

A1. Configure and enable Db2 Text Search Services Support manually

1. Backup and empty the content of the constructor.xml file that is inserted with the Db2TS Support installation.
2. For IBM Db2 Text Search Services, copy everything between the begin and end tag from the appropriate document source sample configuration file in the samples sub directory of the installation folder where Db2 Text Search Services Support is installed and add this to the \config\constructors.xml file between the begin and end tag. The configuration file must contain separate sections for each document source that you selected at installation time.
3. Add a password file called afu_ctor_password.dat in the <Db2TextSearchServicesInstallDir>\plugins\ directory of IBM Db2 Text Search Services and add the password in plain text format. This file can be configured using the <ArchivePasswordFile> tag. The password is encrypted during the initialization of Db2 Text Search Services Support when IBM Db2 Text Search Services is started.
4. Adapt constructor configuration settings as per the samples provided. The attribute mappings section includes custom mappings that you might want to modify.

5. The trace and log settings for Db2 Text Search Services Support are set to false by default. Modify the trace level and the dump settings as per your need.
6. You should not remove any of the attributes and archive attributes that are in the configuration sample files because they are required by Content Collector. Only add new entries to the existing attributes or adapt the existing attributes if required.
7. Start the Db2 Text Search Service. Verify the service is started successfully.

A2. Creating a manual catalog prior to initial configuration run

1. Complete the upgradation or fresh installation of ICC FP13.
2. Create a manual catalog and make this catalog active.
3. Navigate to Connectors → IBM Content Manager → CM8.x Connections.
4. Create a new CM8 connection providing the Username, password, and repository details. Save the connection details.
5. Navigate to Full Text Search Services Support.
In the center pane, select Support for Email (Select the appropriate Source required)
In the right pane, under General tab, select the CM8 connection as the repository connection. Save the repository connection details.
6. If there were multiple sources selected during Db2 Text Search Support installation on the CM8 server, then the repository connection needs to be selected for all those corresponding sources in Configuration Manager,
For example – If the Db2 Text Search Support was installed selecting Email and SharePoint sources, then in the Configuration Manager, the CM8 connection should be selected as the repository connection under both Support for Email and Support for SharePoint.
7. During Manual catalog creation, Support for File Systems source page will not be shown in the Full Text Services Support. Follow the below step to select repository connection in Web Application and then navigate back to the Full Text Search Services Support.
8. Navigate to General Settings → Web Application → General → Repository Connection
Click on Add and select the CM8 connection added as the Repository name in the dropdown. Save the repository connection details.
9. For Selecting the repository connection for File Systems, navigate now to Full Text Services Support and select the repository connection and save the details.
10. Restart the ICC Web Application Liberty service. Verify the service is started successfully. Run the below Remote Setting URL in the browser for the appropriate ConfigurationKey based on the source.

Email:

https://<ICC_hostName>:11443/AFUConfig/Configuration?type=ibm.ctms.ContentSearchService&unique=Email

SharePoint:

https://<ICC_hostName>:11443/AFUConfig/Configuration?type=ibm.ctms.ContentSearchService&unique=SharePoint

File Systems:

https://<ICC_hostname>:11443/AFUConfig/Configuration?type=ibm.ctms.ContentSearchService&unique=FileSystem

Ensure that the below tags are a part of the response returned on hitting the URL in the browser.

```
<cm8Config definitionId="<CM8 Connection Id">">
  <archiveType>ICM</archiveType>
  <schema><schema> </schema>
  <databaseName><database> </databaseName>
  <credential user="<UserName>" password="<password>"/>
</cm8Config>
```

11. Restart the Db2 Text Search Service. Verify the service is started successfully.

B. Configuring the repository in existing NSE catalog

1. Complete the upgradation to ICC FP13.
2. In the NSE catalog which was active, navigate to General Settings → Full Text Search Services Support
In the center pane, select Support for Email (Select the appropriate Source required)
In the right pane, under General tab, select the CM8 connection as the repository connection.
Save the repository connection details.
3. If there were multiple sources selected during Db2 Text Search Support installation on the CM8 server, then the repository connection needs to be selected for all those corresponding sources in Configuration Manager,
For example – If the Db2 Text Search Support was installed selecting Email and SharePoint sources, then in the Configuration Manager, the CM8 connection should be selected as the repository connection under both Support for Email and Support for SharePoint.
4. Restart the ICC Web Application Liberty service. Verify the service is started successfully.
Run the below Remote Setting URL in the browser for the appropriate ConfigurationKey based on the source.

Email:

https://<ICC_hostname>:11443/AFUConfig/Configuration?type=ibm.ctms.ContentSearchService&unique=Email

SharePoint:

https://<ICC_hostname>:11443/AFUConfig/Configuration?type=ibm.ctms.ContentSearchService&unique=SharePoint

File Systems:

https://<ICC_hostname>:11443/AFUConfig/Configuration?type=ibm.ctms.ContentSearchService&unique=FileSystem

Ensure that the below tags are a part of the response returned on hitting the URL in the browser.

```
<cm8Config definitionId="<CM8 Connection Id">">
  <archiveType>ICM</archiveType>
  <schema><schema> </schema>
  <databaseName><database> </databaseName>
```



```
<credential user="<UserName>" password="<password>"/>
</cm8Config>
```

5. Restart the Db2 Text Search Service. Verify the service is started successfully.

Item Type creation

In case the customer had proceeded with option 1b in the above section, then the manual catalog can be completed to create a full catalog configuring the source connection, archived data access for searching, etc.

OR

Run through the initial configuration screens to complete catalog creation with Db2TS enabled CM8 repository and once completed, discard the manual catalog that was created.

Once initial configuration creates an item type successfully in CM8 server, perform the below checks of verification.

- In the CM8 server, login to the System Admin Client. A new email item type generated will be available and can be viewed in the System Admin Client on the CM8 server.
- Run the below query to obtain the componenttypeid of the item type containing the documents marked for indexing

```
COMPONENTTYPEID = SELECT COMPONENTTYPEID FROM
ICMADMIN.ICMSTCOMPVIEWDEFS, ICMADMIN.ICMSTNLSKEYWORDS WHERE
KEYWORDNAME = '<ItemTypeName(case-sensitive)>' AND KEYWORDCODE =
COMPONENTTYPEID AND KEYWORDCLASS = 5 FOR READ ONLY WITH UR
```

- Use the following query to find the componentType table
Replace the COMPONENTTYPEID obtained from the above results to the below query:
SELECT * FROM ICMUT0<COMPONENTTYPEID>001 FOR READ ONLY WITH UR
- The presence of above 2 tables associated with the item type, confirms that the item has been successfully inserted in the database.

Log / Dump / Cache files

Log Files

1. trace0.log
 - The Db2TS service start-up logs are logged in to the trace0.log log file in the Db2COPY directory log path.
 - The default path on windows is in <Db2 installDir>\Db2COPY1\Db2\db2tss\log. On non-windows machine, the default directory points to <db2 instanceDir>/sqllib/db2tss/log/.
2. ContentCollector_YYYYMMDD_tttt*.log

- The IBM specific constructor initialization startup logs are logged to the ContentCollector_yyyymmdd_tttt*.log in the <SQLLIB directory>\db2tss\log path on windows. For non-windows, the default path is <db2InstanceDir>\sqllib\db2tss\log/.
 - Only when the Log Level is set to Trace, both the ContentCollector_yyyymmdd_tttt.trace.log and ContentCollector_yyyymmdd_tttt.system.log will be generated. Otherwise, only the ContentCollector_yyyymmdd_tttt.system.log file will get created.
 - Log files use timestamps in their names so that there is a new set of log files each time the Db2 Text Search Services Support run. By default, only one log file is generated for a service start and the indexing logs of all the types of source documents are logged to the same file.
3. AfuIndexerDb2ts*.log
- This log file will be generated only when the indexing is invoked on an item type.
 - Depending on the same log level set the AfuIndexerDb2ts*.log files are generated. Only when the Log Level is set to Trace, both the AfuIndexerDb2ts.trace.log and AfuIndexerDb2ts.system.log will be generated. Otherwise, only the AfuIndexerDb2ts.system.log will get created.
 - The AfuIndexerDb2ts*.log files are written to the SQLLIB\db2tss in Windows and to the db2inst1\sqllib\db2tss directory in case of non-windows. If the above paths are not found, then the logs are created under < Db2TextSearchServicesSupportInstallDir>\log directory.
 - The content of the logs file are replaced with every indexing run.

Dump Files

- When the dumps are enabled, as per the dump selection, doc/txt/xml sub directories are created inside the log/dump directory
- Note that it is recommended to set the dumps to false in constructor.xml for performance reasons.

Cache Files

- On successful initialization of the configured Email / SharePoint constructor with Db2TS server start, the corresponding .cache file will be generated inside plugins directory.
 - afu_config_Email.cache
 - afu_config_SharePoint.cache
- For File Systems, the .cache file will be generated when the indexing is invoked on any item type post the service start.
 - afu_config_FileSystem.cache

Once the Db2TS service is started,

1. Ensure that the Db2TS server logs of trace0.log file does not show any errors.
2. Ensure that the ICC specific log files of ContentCollector_*.log on successful start of the Db2TS service log does not show any errors.

3. If the dumps are enabled, the dump directory gets created in the log location along with the sub-directories for each of the dumping options of doc, txt, xml.

Remote config changes Config Manager UI changes

The following changes will be seen with the latest FP of ICC to the catalog created with Text search.

This section contains the details of log settings, indexing attributes, etc. that are required for indexing. The details of which are fetched by the Remote Settings URL which gets configured in the constructors.xml file upon Db2TS Support installation.

Navigate to the General Settings > Full Text Search Services Support.

1. When an initial configuration catalog will be created for any source with CM8 repository, Repository Connection dropdown should be auto filled with CM8 connection.
2. New setting will be present for "File System Source" only for CM8 repository and will be created when a new catalog is created with initial configuration by selecting source as "File System Source" and repository as CM8. UI will not allow making any changes to existing created definition, so all sections in the right panel will be read-only for "File System Source" definition:
All the sections in the right panel will be editable for all other sources except "File System Source".
3. When any change is made to the Full Text Search Services Support section, on click of save a dialog box will open, on click of which restart server message will get displayed. The message indicates the need to restart the Db2TS service for the changes to get reflected.

This section also includes attribute settings for documents to be indexed including settings for the repository connection and the dump facility.

Common log settings for Full Text Search services support

The log and trace options for IBM Content Collector Db2 Text Search Services Support are common to all the configured source document pre-processors and when enabled apply to all documents that are pre-processed by Db2 Text Search services support.

Because logging and tracing cannot be enabled separately for email, or Microsoft SharePoint documents, the log files can contain a mix of all source documents and can become very large.

The log settings include these settings:

The default level is Error. Log entries are cumulative; each level is inclusive of the levels preceding it in the list of available log levels. A log level of type Information, for example, will include Trace, Information, Warning, and Error log entries.

A log file location which specifies the logging directory to which the Db2 Text Search Services Support log and trace files are written. The default directory is .log, the log subdirectory where IBM Db2 Text Search Services is installed and contains the bin directory containing service start up script files.

Note: The log directory must be a directory on the server on which IBM Db2 Text Search Service runs and not the server on which the IBM Content Collector Configuration Manager runs.

General configuration settings for all source documents in Full Text Search services support

The general configuration settings for documents that are pre-processed by Db2 Text Search Services Support include settings for the repository connection and the dump facility.

On the General tab for each document type in the IBM Content Collector Configuration Manager under General Settings > Full Text Search Services Support, you can define the following settings:

Repository connection: Specifies the access information for the Content Manager. You can select the repository connection that provides the database and login properties for Content Manager Connection.

Dump settings:

- Create dump files for all input documents: equivalent to enabling **DumpDocInput** in the constructors.xml file.
- Create dump files of the XML file for indexing: equivalent to enabling **DumpXmlOutput** in the constructors.xml file.
- Create dump files for the textual content of attachments: equivalent to enabling **DumpTxtOutput** in the constructors.xml file.

Dump directory: specifies the name of the directory to which dump files are written in corresponding subdirectories. The default directory is `./log/dump`, the log subdirectory where IBM Db2 Text Search Services is installed and contains the bin directory containing service start up script files.

On Windows, it will be pointing to the db2 installation directory. The default path is pointing to `\\sqllib\db2tss\log\dump` directory.

On non-windows, it will be pointing to the log directory under db2 instance user. The default path is pointing to `/db2inst1/sqllib/db2tss/log/dump` directory.

System and custom configuration settings for email documents:

The configuration for email documents includes default attribute and archive attribute settings, and custom settings that can be added by the user.

Adding custom attributes

Add custom attributes, custom archive attributes, and Full Text Search Services Support configuration option settings on the Custom Settings tab.

The custom attributes and custom archive attributes must be added under the respective sections in the format shown in the example.

Note the following when adding attributes:

System and custom configuration settings for File Systems documents

The configuration for File Systems documents does not include any user defined attributes. The default attribute and archive attribute settings, and custom settings section has been included and displayed on the Full Text Search Services Support. However, the [Attributes], [ArchiveAttributes], [Settings] are left blank and disabled.

There are also no default archive attributes for File Systems documents that are required by Content Collector.

Note: Any change made to the Full Text Search Services Support section requires the Db2Text Search service to be restarted for the changes to get reflected.

Setup Tools

CM Repository configuration

This tool can be used has to create a fresh item type to the CM8 repository without the need to create a new catalog by running through initial configuration.

CM Repository enablement

The support for Text Search option available with this tool has been removed. The Text Search Enablement suing this tool will no longer be supported from FP13.

Indexing and Reindexing

Once the catalog has been created and the items are archived, we can then index them at any later point of time, with the help of the `afulIndexer` command.

`afulIndexer` usage

```
-it <itemtype>  
[-config <configFile>]  
[-name <configuration constructor name>]
```

`-config configFile`

The file path name of the IBM Db2 Text Search Services configuration file that contains the Db2 Text Search Services Support configuration, Ex -
<Db2TextSearchServicesInstallDir>\config\constructors.xml.

`-name constructorName`

The name of the constructor plug-in that is used by IBM Db2 Text Search Services for pre-processing the source documents; AFUCTOR4EMAIL for email documents or AFUCTOR4SHAREPOINT for Microsoft SharePoint documents and ICMDCTOR for File Systems documents. The name that you enter must match the constructor's name that is used in the constructors.xml configuration file.

-it <itemtype>

The name of the item type containing the archived items to be indexed

Indexer command

To execute indexing for Email or SharePoint or File Systems, on CM8 server navigate to the <Db2TextSearchServicesSupportInstallDir> to run the afulIndexer command.

PF8 the example commands for indexing on Windows platform:

EMAIL

```
afulIndexer.cmd -it ICCSMTPCmp -config  
C:\ProgramData\IBM\Db2\Db2COPY1\Db2\db2tss\config\constructors.xml -name  
AFUCTOR4EMAIL
```

```
afulIndexer.cmd -it ICCEmailCmpEX -config  
C:\ProgramData\IBM\Db2\Db2COPY1\Db2\db2tss\config\constructors.xml -name  
AFUCTOR4EMAIL
```

```
afulIndexer.cmd -it ICCEmailCmpLD -config  
C:\ProgramData\IBM\Db2\Db2COPY1\Db2\db2tss\config\constructors.xml -name  
AFUCTOR4EMAIL
```

SHAREPOINT

```
afulIndexer.cmd -it ICCSharepointDM -config  
C:\ProgramData\IBM\Db2\Db2COPY1\Db2\db2tss\config\constructors.xml -name  
AFUCTOR4SHAREPOINT
```

FILE SYSTEMS

```
afulIndexer.cmd -it ICCFilesystem -config  
C:\ProgramData\IBM\Db2\Db2COPY1\Db2\db2tss\config\constructors.xml -name ICMDCTOR
```

Indexing execution

When the document is marked for indexing, the attributeld column associated with the attribute of ICCIdxRCString is populated with the error code or success codes. For more details, refer to Error Codes section.

Reindexing

Reindex command for Windows platform:

```
afulIndexer.cmd -it <itemType> -config <\config\constructors.xml> -name <AFUCTOR4EMAIL/  
AFUCTOR4SHAREPOINT / ICMDCTOR > <reindex options>
```

Reindexing options

- `reindexrc <idxrc>` (reindex using the idxrc error code)
Reindexes all documents that were indexed during a previous indexing run and that produced a specific value.
- `reindexerrors` (reindex all docs that have errors)
Reindexes all documents that were indexed during a previous indexing run and that resulted in an error IDXRC
- `reindexwarnings` (reindex all docs that has warnings)
Reindexes all documents that were indexed during a previous indexing run and that resulted in a warning IDXRC
- `reindexitems <itemid>;<itemid>...`(reindex the list of Itemids provided)
The itemIds should be separated by a : when executing the command on non-windows platform.
Reindexes only those documents that are listed
- `reindexdate <yyyymmdd>` (reindex all docs archived after the date)
Reindexes all documents that have date of <yyyymmdd> or later.
- `reindexsearch <term>`
Reindexes those documents that were indexed during a previous indexing run and where the specified string was indexed with the document.
- `reindexsearch <Query to search>`

Example:

```
-reindexsearch (@xmlxp:"//icc_subject[.contains(\"january\")])  
-reindexsearch "(@xmlxp:"//icc_attachment[.contains(\"january is cold\")])"  
- reindexsearchfile <*filename>
```

Example:

```
[reindexsearch] APAR1234 = @xmlxp:"//icc_subject[.contains(\"january\")]"APAR5678 =  
"wrong"
```

* filename : It is an ini file that can have multiple search terms, the Key configured and the name of the file can be any custom name. Its same as NSE.

Note: The log files to be checked for traces remain same as that of Indexing.

Indexing log files

In cases of any issues during indexing, the logger of the indexing can be verified in the below two files:

1. `AfulIndexerDb2ts*.log`
This log file contains the trace/error occurred during the connection to CM8 or while updating the error codes.

2. ContentCollector_YYYYMMDD_tttt*.log
The log file contains the trace/error occurred during the pre-processing of the documents marked for indexing

Verifying the generated indexing code

To identify document-related indexing failures:

1. Run the following queries, to obtain the attribute Id associated with the 'ICCIDxRCString'
 - a. Obtain the keywordcode of the item type for the attribute of 'ICCIDxRCString' by replacing the COMPONENTTYPEID obtained from the Item Type Creation section.

```
KEYWORDCODE = SELECT KEYWORDCODE FROM ICMSTNLSKEYWORDS,
ICMSTCOMPVIEWATTRS WHERE KEYWORDCLASS = 1 AND KEYWORDNAME
= 'ICCIDxRCString' AND COMPONENTVIEWID = '< COMPONENTTYPEID>' AND
ATTRIBUTEID = KEYWORDCODE
```

```
IDXRC = "ATTR000000"+ KEYWORDCODE
```

2. The IDXRC attributeId column is a part of the component type table associated with the item type. Use the following query to find the componentType table

Replace the COMPONENTTYPEID obtained from 1a) to the below query:

```
SELECT * FROM ICMUT0< COMPONENTTYPEID >001 FOR READ ONLY WITH UR
```

Look for the itemID logged on the console to identify the item indexed. Run the above query to obtain the results from the component table. The error/success code can be found against this itemID in the column name obtained from 1b.

The table will contain the error codes are as code+E and warnings will be written as code+W.

Example: If a warning of 32 is thrown, the code will be written as 32W. If an error of 8 is thrown, the code will be written as 8E. In case there are multiple errors thrown from a single document, then the codes are written with comma separated values as 32W, 8E.

Error codes

Expected Error codes:

IDXRC	ERROR STRING	ERROR MSG STRING
0	AFU_INDEXING_SUCCESSFUL	MSG_AFU_INDEXING_SUCCESSFUL
1	AFU_INDEXING_GENERAL_WARNING	MSG_AFU_INDEXING_GENERAL_WARNING
2	AFU_CNV_DISABLED	MSG_AFU_CONVERSION_DISABLED
4	AFU_DOC_ENCRYPTED	AFU_DOC_ENCRYPTED
8	IQQP0031E	MSG_PASSWORD_PROTECTED_DOCUMENT
16	IQQP0026E	MSG_UNRECOGNIZED_ENCODING
32	IQQP0027W	MSG_NO_TEXT_EXTRACTED
64	IQQP0008W	MSG_UNSUPPORTED_TYPE
128	IQQP0032E	MSG_MALFORMED_DOCUMENT
256	AFU_DOC_PARTIAL	MSG_AFU_DOCUMENT_PARTIAL
512	IQQP0012W	MSG_TRUNCATED_DOCUMENT
1024	IQQP0033W	MSG_FALLBACK_PARSER_INVOKED

2048	IQQP0051E	MSG_UNRECOGNIZED_LANGUAGE
3355443 2	AFU_INDEXING_GENERAL_ERRO R	MSG_AFU_INDEXING_GENERAL_ERRO R
3355443 3	AFU_DOC_ERROR	MSG_AFU_DOCUMENT_GENERAL_ERR OR
3355443 4	AFU_INTERNAL_PARSER_ERRO R	MSG_AFU_INTERNAL_PARSER_ERROR
3355443 5	IQQP0052E	MSG_INVALID_ARCHIVE_CONTENT
3355443 6	IQQP0028E	MSG_MALFORMED_XML_DOCUMENT
3355443 8	IQQP0013E	MSG_SIZE_LIMIT_EXCEEDED

If there are any errors thrown besides the errors listed in the above table, it will be considered as MSG_AFU_INDEXING_GENERAL_ERROR

If there are any warnings thrown besides the warnings listed in the above table, it will be considered as MSG_AFU_INDEXING_GENERAL_WARNING

Search

If the catalog is created through initial configurations, the db2ts config file and mapping file configurations are imported automatically to the to the Archived Data Access and no manual changes are required.

In case the catalog is created manually, then the below templates need to be modified to provide the details of the item type to be searched and imported to the Archived Data Access.

1. afu_db2ts_search_config_template.xml
2. afu_cm8_compound_db2ts_search_mapping_template.xml

The mapping file will be required to be updated with the item type and its corresponding EChild and AChild to perform search operation.

Note: The IBM Web Application Liberty server needs to be restarted post the file imports.

Federated search

Once we have moved to the Db2TS indexes and started archiving to the newly created item type. To support searching of the documents against both NSE (already indexed docs prior to using Db2TS) and Db2TS index, we need to modify the mapping templates imported to the Archived Data Access section.

This can be achieved either by updating the configurational files manually or through the configuration manager UI.

Changes to Archive Mapping template file

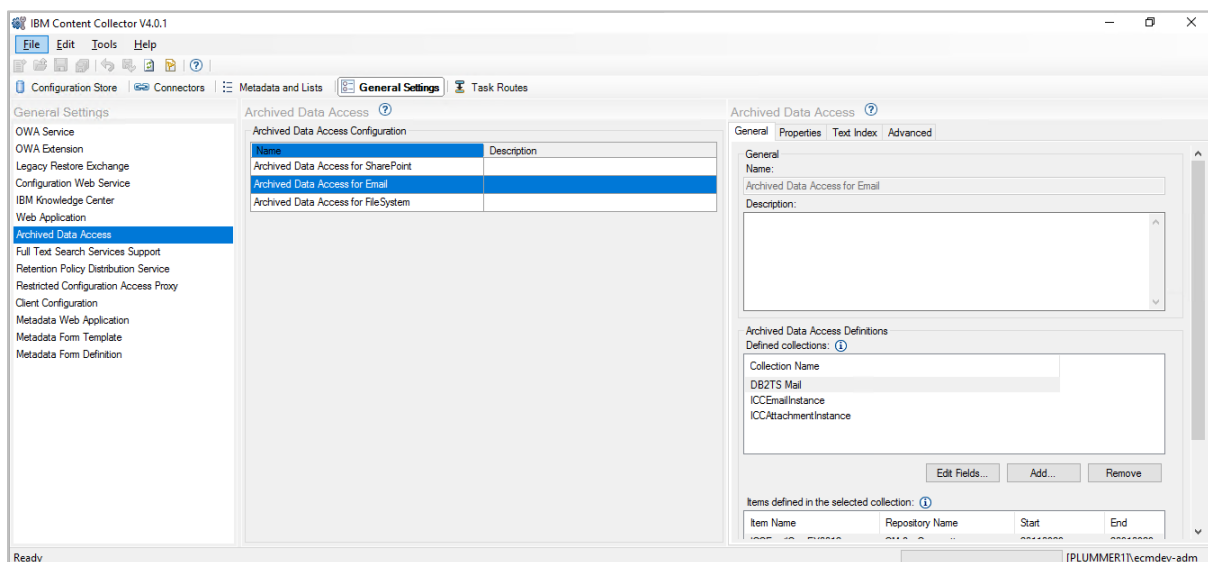
Modifying the Archive Mapping template through Configuration Manager UI

With the earlier versions of ICC, the Archived Data Access section of Configuration Manager provides the options to configure the mapping files for searching against multiple item types. With this release of ICC, this feature has been leveraged to allow the configuration of NSE (already indexed) items along with Db2TS item types for federated search.

- The item type of Db2TS is added with a collection name of “Db2TS Mail”.
- The item type of NSE is added with a collection name of “NSE Mail”

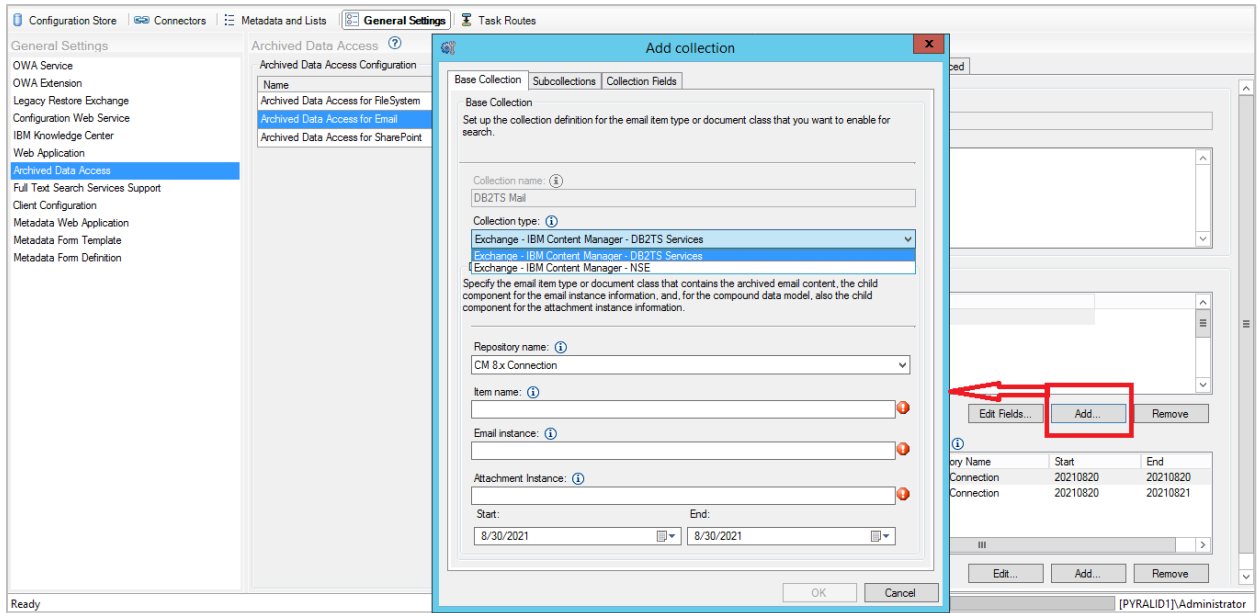
To configure the mapping file, navigate to General Settings >> Archived Data Access

1. After completing initial configuration, when Archived Data Access Configuration section is visited, Definition collections list displays Db2TS Mail and its related instances for Db2TS.

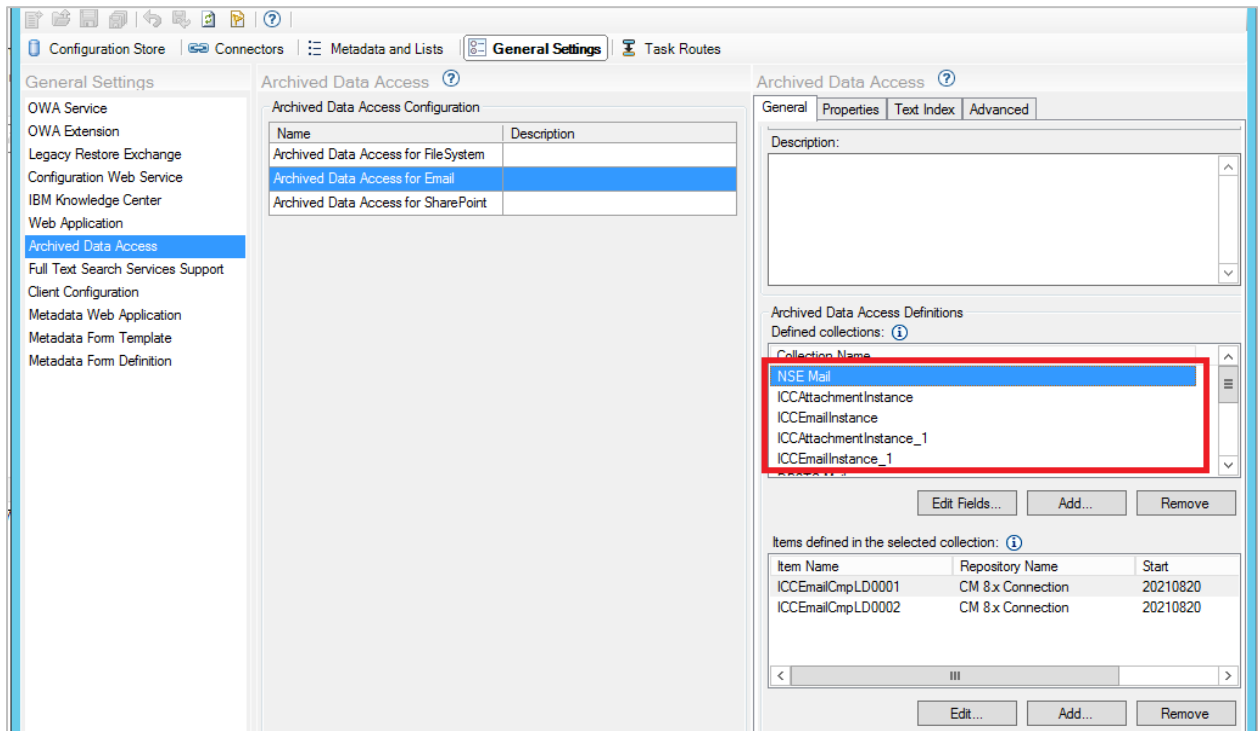


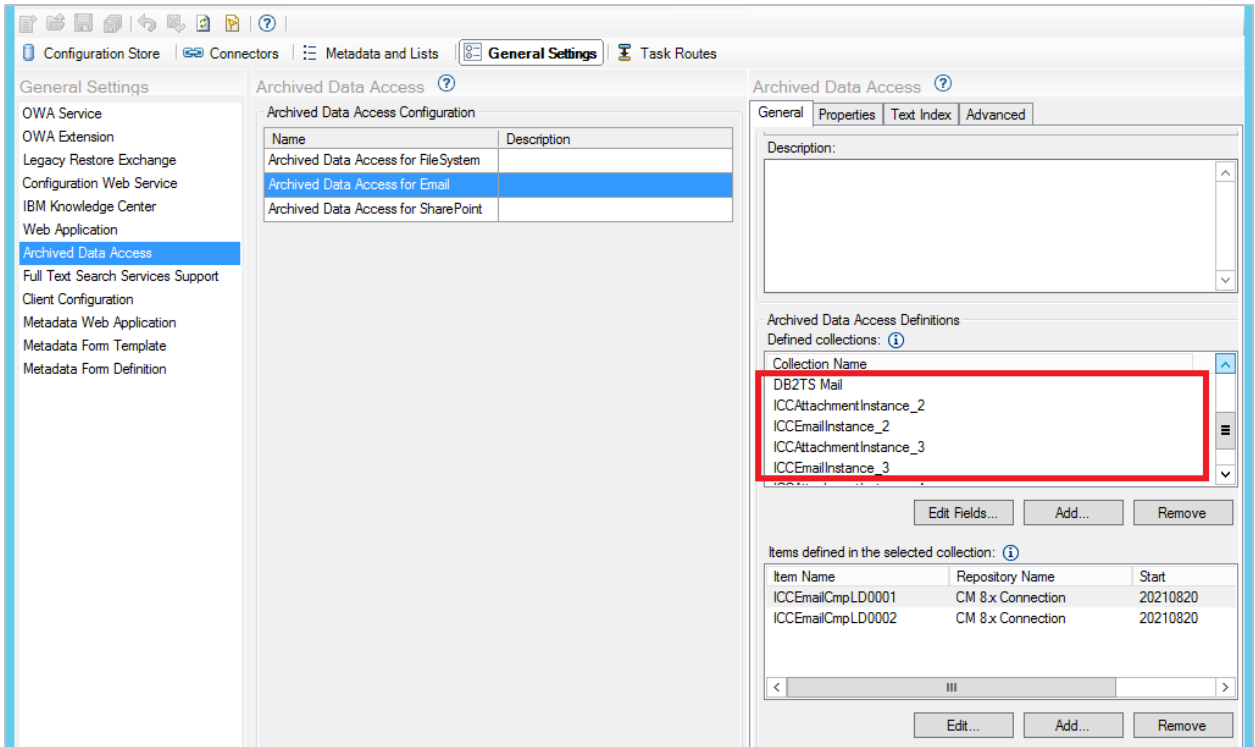
2. In Archived Data Access Configuration, when email source is selected, and then “Add” button is clicked from Defined Collections list area; a popup gets opened to Add Collection.

With this release, a new dropdown has been added for Collection type containing the values for Db2TS or NSE. Depending upon the values chosen in the dropdown, the Collection Name will be defaulted to read-only prefilled value of “Db2TS Mail” or “NSE Mail”.

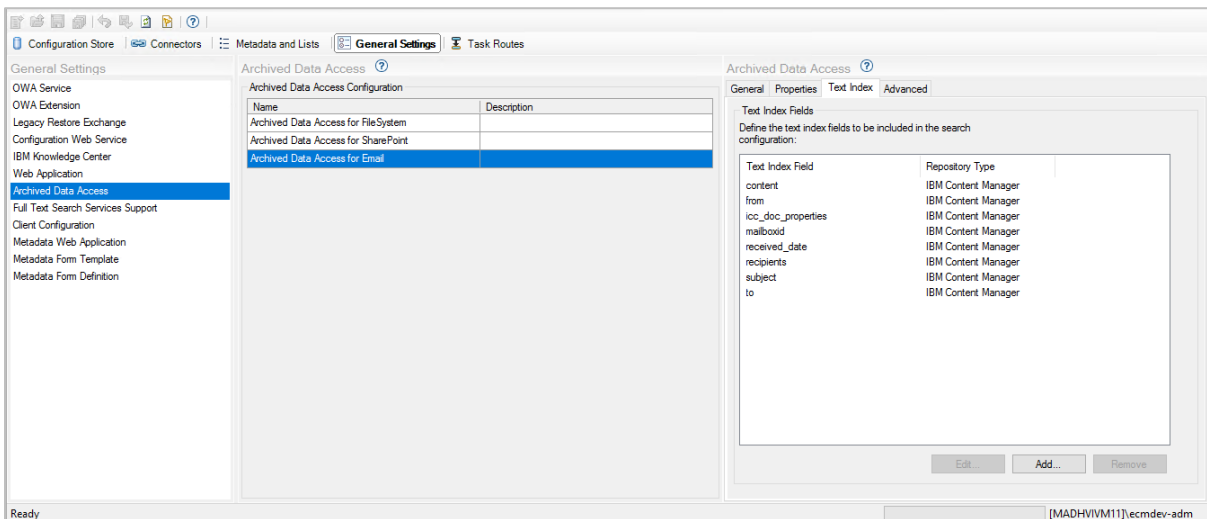


- After adding new Collection Type and clicking “OK” in above shown popup, then new document type, related email & attachment instance will be created inside existing collection “NSE Mail” or “Db2TS Mail”.
If already collection type instance exists for “NSE Mail” or “Db2TS Mail”, then new email and attachment instance will be added to existing collection type instance.

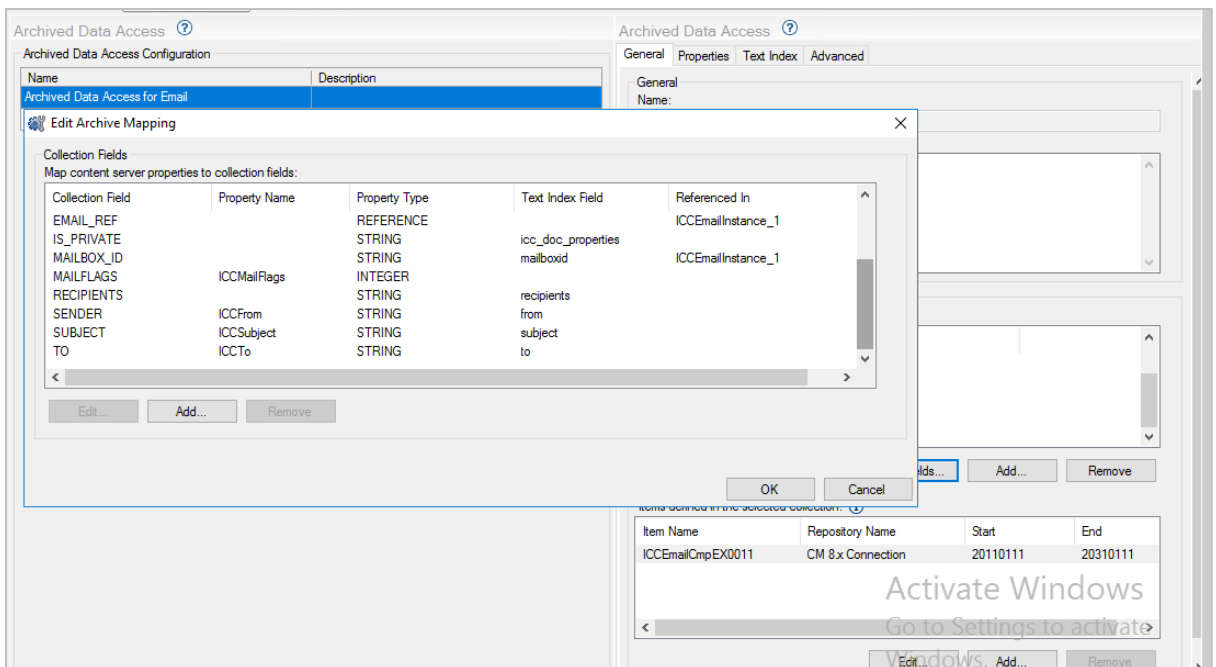
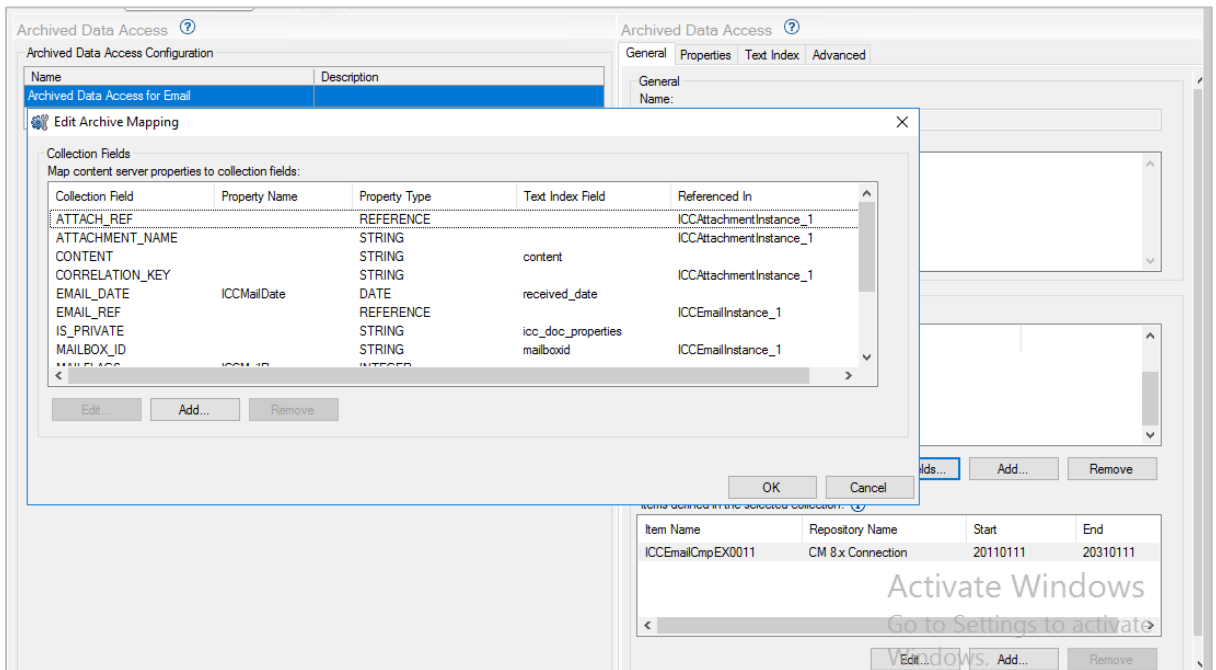




- Once an item type is added from the General tab, we need to add the Text Index fields from the Text Index tab. For an NSE Mail type, all the listed fields in the screenshot need to be added to Text Index fields by clicking on Add button.



- Navigate back to the General tab, select the Collection Field of NSE Mail and click on Edit Fields. The Text Indexes added need to be associated with the item type from this window by clicking on Add button for all the below fields.



Modifying the Archive Mapping template manually

Alternatively, the mapping file modifications can also be done manually. With db2ts, template file of "afu_cm8_compound_db2ts_search_mapping_template.xml" has been introduced and deployed in the ICCInstallDir>\Configuration\initialConfig\data\search location

This file contains the mapping file details of Db2TS. To configure the files for NSE, append the below section to the above mapping file.

```
<doc_type_collection id="NSE Mail" nm="NSE Mail" collectionType="%ICC_EMAIL_PROVIDER%">
  <repositories>
    <repository id="empty">
      <doc_types>
        <doc_type>
          <name>%ICC_CONFIG_ITEMTYPE_NAME%</name>
          <children>
            <child
ref_coll="ICCEmailInstance">%ICC_CONFIG_ITEMTYPECHILD_EI%</child>
            <child
ref_coll="ICCAttachmentInstance">%ICC_CONFIG_ITEMTYPECHILD_AI%</child>
          </children>
          <start>%ICC_SEARCHDATE_MIN%</start>
          <end>%ICC_SEARCHDATE_MAX%</end>
        </doc_type>
      </doc_types>
    </repository>
  </repositories>

  <fields>
    <field nm="EMAIL_DATE" type="DATE" partitionkey="true">
      <attr>ICCMailDate</attr>
      <search format="yyyyMMddHH">received_date</search>
    </field>
    <field nm="SENDER" type="STRING">
      <search>from</search>
      <attr>ICCFrom</attr>
    </field>
    <field nm="RECIPIENTS" type="STRING" multivalue="true">
      <search>recipients</search>
    </field>
    <field nm="TO" type="STRING">
      <search>to</search>
      <attr>ICCTo</attr>
    </field>
    <field nm="SUBJECT" type="STRING">
      <search>subject</search>
      <attr>ICCSubject</attr>
    </field>
    <field nm="CONTENT" type="STRING">
      <search>content</search>
    </field>
    <field nm="IS_PRIVATE" type="STRING">
      <search>icc_doc_properties</search>
    </field>
    <field nm="MAILFLAGS" type="INTEGER">
      <attr>ICCMailFlags</attr>
    </field>
    <field nm="MAILBOX_ID" type="STRING">
      <search>mailboxid</search>
      <reference>EMAIL_REF.MAILBOX_ID</reference>
    </field>
    <field nm="EMAIL_REF" type="REFERENCE" ref_coll="ICCEmailInstance_1" multivalue="true">
```

```

    <relationship type="CHILD"></relationship>
  </field>
  <field nm="ATTACH_REF" type="REFERENCE" ref_coll="ICCAttachmentInstance_1"
  multivalue="true">
    <relationship type="CHILD"></relationship>
  </field>
  <field nm="CORRELATION_KEY" type="STRING">
    <reference>ATTACH_REF.CORRELATION_KEY</reference>
  </field>
  <field nm="ATTACHMENT_NAME" type="STRING">
    <reference>ATTACH_REF.ATTACHMENT_NAME</reference>
  </field>
</fields>
</doc_type_collection>

```

When configuring the files manually, replace all the place holders with appropriate values. The placeholder of %ICC_EMAIL_PROVIDER% should be replaced with the collection type of NSE/Db2TS as per the section being modified.

Below collection types have been introduced for Exchange/Domino

1. ICC_Exchange_Email_Db2TS_Compound
2. ICC_Domino_Email_Db2TS_Compound

Below existing collection types were used with Exchange/Domino (NSE)

1. ICC_Exchange_Email_Compound
2. ICC_Domino_Email_Compound

Please make sure that the doc_type_collection tag is present with an id of "NSE Mail" for NSE section and with an id of "Db2TS Mail" for Db2TS section. The corresponding EChild and AChild instance must be also configured against the NSE item type in the mapping file.

Please note this EChild and AChild instance must have an id/nm different that the child instances added for db2ts. Hence, we have provided the id/nm appended with an "_1" and the appropriate changes are also made to the item type section fields referring to these child elements with ref_coll.

```

<doc_type_collection id=" ICCEmailInstance_1" nm=" ICCEmailInstance_1" collectionType="CMCHILDCOMP">
  <repositories>
    <repository id="empty">
      <doc_types>
        <doc_type>
          <name>%ICC_CONFIG_ITEMTYPECHILD_EI%</name>
        </doc_type>
      </doc_types>
    </repository>
  </repositories>
  <fields>
    <field nm="MAILBOX_ID" type="STRING">
      <attr>%ICC_CONFIG_ATTR_MAILBOXID%</attr>
    </field>
  </fields>
</doc_type_collection>

```

```

<doc_type_collection id="ICCAttachmentInstance_1" nm="ICCAttachmentInstance_1"
collectionType="CMCHILDCOMP">
  <repositories>
    <repository id="empty">
      <doc_types>
        <doc_type>

```

```

                <name>%|ICC_CONFIG_ITEMTYPECHILD_AI%</name>
                </doc_type>
            </doc_types>
        </repository>
    </repositories>
    <fields>
        <field nm="ATTACHMENT_NAME" type="STRING">
            <attr>%AFU_CONFIG_ATTR_FILENAME%</attr>
        </field>
        <field nm="CORRELATION_KEY" type="STRING">
            <attr>%|ICC_CONFIG_ATTR_CORRELATIONKEY%</attr>
        </field>
        <field nm="CONTENT_REF" type="REFERENCE" ref_coll="EDMDefaultContent">
            <attr>%|ICC_CONFIG_ATTR_CONTENTREF%</attr>
        </field>
        <field nm="CONTENT" type="STRING">
            <reference>CONTENT_REF.CONTENT</reference>
        </field>
    </fields>
</doc_type_collection>

```

Changes to Archive Config file

Once the Archive Mapping file has been updated, the corresponding changes to the config file must be made manually. With db2ts, template file of "afu_db2ts_search_config_template.xml" has been introduced and deployed in the ICCInstallDir>\Configuration\initialConfig\data\search location

Open the config template file containing the changes for Db2TS and modify the file to uncomment the commented section of "NSE Mail"

Navigate to the Configuration Manager UI and import the modified config template to the Archived Data Access section and restart the IBM Web Application Server for the changes to reflect.

PFA the sample configuration files templates for federated search consisting of both the NSE and Db2TS item types. These have been provided for reference purpose and need to be modified with the appropriate item type names.



Sample federated
search templates.zip

To enable search by folder name

1. Make the information available in the repository.
 - a. Create one new attribute with name ICCMailFolder (This can be any custom name). The property must be of the Variable Character Attribute type and Extended Alphanumeric Character type. Set the maximum Character length that is allowed for a property value to 1024.
 - b. Add this attribute to the child component AFUEChild of email item type.
 - c. To make the property available for mapping in a task route, restart Configuration Manager.

- d. If your source system is Lotus® Notes®, enable the Processing option of “Provide folder information” in the Email Connector configuration.
 - e. Adapt your email archiving task route to have the metadata archived.
 - f. On the Item Type Configuration tab of the CM8 Configure Item Types task, map <Email,Folder> to the folder information property that you defined on the AFUEChild.
 - g. To enable search on special mailbox folders in language-independent format, map <Email, Folder(normalized)> to the folder information property. In this case, strings like %INBOX% are stored in the repository. You can then define custom names for these special folders by which users can search.
 - h. In Configuration Manager, navigate to the Custom Settings tab of the Full Text Search Services Support configuration for email documents.
Add the following options to the [ArchiveAttributes] section:
ICCMailFolder = /icc_document/icc_mailfolder
Ensure the attribute path contains the custom attribute name created to store the folder name.
 - i. ICC Search related configuration also needs to be updated to enable the search against the folder name.
2. Add the definitions to the mapping file for archived data access.
 - You must add the definitions manually to the archive mapping file.
 - Export the existing mapping file from the configuration store to disk.
 - a. In Configuration Manager, select General Settings > Archived Data Access > Archived Data Access for Email > Advanced and export the file.
 - b. Open the search_mapping.xml file with a text editor.
 - c. In the <fields> section of the collection definition that contains the attributes that are common to all email instances, add the following element


```
<field nm="ICCMailFolder" type="STRING">
  <search>/icc_mailfolder</search>
  <reference>EMAIL_REF.ICCMailFolder</reference>
</field>
```
 - d. In the <fields> section of the EChild instance of collection definition add the following element


```
<field>
  <field nm="ICCMailFolder" type="STRING">
  <attr>ICCMailFolder</attr>
</field>
```
 - e. Save the file.
 - f. Import the updated file into the configuration store and save the new configuration.
 3. Adapt the layout of the search page by updating the search configuration file
Update the <declaration> section, the <form> section, and the <result> section.
 - a. Export the search configuration file from the configuration store to disk.
 - b. Open the search_config.xml file with a text editor.

- c. Define the folder name field as a property in the <declaration> section:

```
<property>
  <name>ICCMailFolder</name>
  <nls-key>jsp.searchrequest.column.folder</nls-key>
</property>
```

- d. Add the field as a property to each collection definitions changed in the search mapping template

```
<property name="ICCMailFolder">
  <field>ICCMailFolder</field>
</property>
```

- e. To add a corresponding input field to the search mask, include the appropriate field definition to the <form> section of the search template.

Create a new group section for folder search:

```
<group id="folder" nls-key="folder.group.header" >
  <field input-type="option">
    <name>ICCMailFolder</name>
    <hidden>>false</hidden>
    <size>100</size>
    <max-length>500</max-length>
    <tooltip-nls-key>jsp.searchrequest.folder.tooltip</tooltip-nls-key>
    <example-nls-key>jsp.searchrequest.folder.example</example-nls-key>
  </field>
</group>
```

- f. To have the folder name displayed in the search result list, include the appropriate column definition in the <result> section of the search template

```
<column>
  <name>ICCMailFolder</name>
  <width>35%</width>
  <nls-key>jsp.searchresult.column.folder</nls-key>
  <caption>>true</caption>
</column>
```

4. Define labels for special folders in the mailbox.

Users can then search by these folder names. These labels are also displayed in the selection list for folder search on the search page and in the search result list.

In your customized properties file, define mappings like the following ones:

%INBOX% Inbox

%DRAFTS% Drafts

%OUTBOX% Outgoing Mail

%SENT% Sent Items

%JUNK% Junk

%RSS_FEEDS% RSS Feeds

%TRASH% Deleted Items

%SYNC_ISSUES% Synchronization Issues

- Restart the IBM Content Collector Web Application service for the changes to take effect.

The search page now contains an additional input field for folder search. Users can enter the folder name to search for email that was in a custom folder in the mailbox. In addition, users can search special mailbox folders by using the name that you defined in your customized properties file. All special folders are also in the drop-down list for the folder search field.

- Update the below configuration options to the <ICCInstallDir>\AFUWeb\afu\config\templates\custom_label.properties file:

```
# Custom folder search attribute
jsp.searchrequest.folder          ICCMailFolder
jsp.searchrequest.folder.tooltip  Enter folder name
jsp.searchrequest.folder.example  Example: &quot;Inbox&quot;

### custom folder search attribute -- column name
jsp.searchrequest.column.folder   ICCFolder
jsp.searchrequest.column.folder.tooltip ICC Folder Name
jsp.searchresult.column.folder    ICC Folder
```

To enable search by attachment name

- Add the definitions to the mapping file for archived data access. You must add the definitions manually to the archive mapping file.
 - Export the existing mapping file from the configuration store to disk.
 - In Configuration Manager, select General Settings > Archived Data Access > Archived Data Access for Email > Advanced and export the file.
 - Open the search_mapping.xml file with a text editor.
 - In the <fields> section of the collection definition that contains the attributes that are common to all email instances, update the following element with the <search> attributes section

```
<field nm="ATTACHMENT_NAME" type="STRING">
    <search>//icc_attachment/@name</search>
    <reference>ATTACH_REF.ATTACHMENT_NAME</reference>
</field>
```

- Save the file.
 - Import the updated file into the configuration store and save the new configuration.
- Adapt the layout of the search page by updating the search configuration file. Update the <declaration> section, the <form> section, and the <result> section.
 - Export the search configuration file from the configuration store to disk.
 - Open the search_config.xml file with a text editor.
 - Define the attachment name field as a property in the <declaration> section:

```

<property>
  <name>ATTACHMENT_NAME</name>
  <nls-key>jsp.searchrequest.attachmentName</nls-key>
</property>

```

- d. Add the field as a property to each collection definitions changed in the search mapping template

```

<property name="ATTACHMENT_NAME">
  <field>ATTACHMENT_NAME</field>
</property>

```

- e. To add a corresponding input field to the search mask, include the appropriate field definition to the <form> section of the search template.

Create a new group section for folder search:

```

<group id="attachmentName" nls-key="attachmentName.group.header">

```

```

  <field>
    <name>ATTACHMENT_NAME</name>
    <hidden>>false</hidden>
    <size>100</size>
    <max-length>500</max-length>
    <tooltip-nls-key>jsp.searchrequest.attachmentName.tooltip</tooltip-nls-key>
    <example-nls-key>jsp.searchrequest.attachmentName.example</example-
nls-key>
  </field>
</group>

```

Unified search

- A new template `afu_cm_compound_db2ts_content_collector_search_mapping` has been added to the below location, for Domino unified search to the below directory:
<ICInstallDir>\AFUWeb\afu\config\templates.
- The below new collection types for Exchange/Domino have been introduced with Db2TS.
 1. ICC_Exchange_Email_Db2TS_Compound
 2. ICC_Domino_Email_Db2TS_Compound
- The template can be updated as per the requirement and imported to the Configuration Manager for search.
- The import option is available on the Advanced tab under General Settings > Archived Data Access > Archived Data Access for Email.

Troubleshooting

With the non-windows CM8 server, during ICC initial configuration, if the configuration on the Content Collector Configuration Steps page fails, with any of the below errors in the Content Manager Library server logs:

```
sqlcode -20427
sqlerrmc <CIE00330 Invalid value " " for parameter "serverport". ". SQLSTATE=3 >
sqlerrp <SQL11013>
```

OR

```
sqlcode -20427
sqlerrmc <"IQQD0040E The client specified the wrong authentication token". SQL >
sqlerrp <SQL11010>
```

Then the below Db2 technote can be followed for resolution:

- <https://www.ibm.com/support/pages/resolving-cie00330-invalid-value-parameter-serverport-sqlcode-20427-sqlstate38h14>
- <https://www.ibm.com/support/pages/db2ts-start-text-gets-iqqd0040e-client-specified-wrong-authentication-token>

Document change history

Date	Description of change
17 December 2021	Original version published