Optim Search
For Data Explorer
Configuration Guide

July 2014

TABLE OF CONTENTS

INTRODUCTION

This document explains how to configure InfoSphere Optim Data Explorer for use with Optim archived data.   When complete, the configuration will result in a Data Explorer project called Optim Search, which will provide the foundation for adding configurations specific to your archive search needs.

Note the instructions contained within this document were created using a Windows platform.   Other platforms will have similar steps.

Prerequisites

This document assumes certain knowledge and prerequisites as follows.

InfoSphere Optim Data Growth Solution

Users of this document are assumed to have working knowledge of InfoSphere Optim Data Growth Solutions, including configuration and use of Open Data Manager (ODM).   This document requires that an ODM data source has already been configured and ready for use.

InfoSphere Data Explorer

Users of this document are assumed to have a base working knowledge of InfoSphere Data Explorer.   It is assumed that Data Explorer v8.2 or greater is installed and ready for use.

Overview

Optim search is a Data Explorer project that allows you to index, crawl and search Optim archived data quickly and easily.   Search results will display the Archive file name for each search result, allowing you to easily identify where data of interest resides within your archive file population.

This document will guide you through the following steps:

1. Installing Optim Search into Data Explorer
2. Making a copy of the Optim Search project (Optional but recommended)
3. Modifying the configuration for your specific environment
4. Indexing, crawling and searching data

INSTALLING OPTIM SEARCH

Optim Search is a Data Explorer project which is a preconfigured project containing components for indexing archived data from InfoSphere Optim Open Data Manager (ODM).   This configuration can be used as a template when adding specific configurations for your environment.   This section explains how to install the Optim Search project and related artifacts into an existing Data Explorer installation.

All artifacts required to complete the configuration are contained in a ZIP file which is located in the following location:

Windows: <OPTIM_INSTALL_ROOT>\rt\search\optim.search.data.explorer_1.0.1.zip

Linux/Unix: <OPTIM_INSTALL_ROOT>/rt/search/optim.search.data.explorer_1.0.1.zip

The contents of the zip files are as follows:

File Name	Purpose
Optim_Search_v1.0.xml	Optim Search import file for Data Explorer
OptimConnect_JDBC_5.3.3.5.zip	Optim Connect JDBC driver
optimSearch.jpg	Optim Search image
Optim_Search_v1.0.1_Configuration.pdf	This document

Copy files needed for Optim Search

This section requires that Data Explorer has been installed and is operational.   If you do not have Data Explorer installed, refer to the Data Explorer installation manual and complete the installation.

Take the zip file "*Data Explorer - Optim Search v1.0.zip*" and decompress the file to a temporary location.   When completed, copy the JPG and JDBC driver as follows:

OptimConnect_JDBC_5.3.3.5.zip   (JDBC Driver)
Unzip the file to *<DATA_EXPLORER_INSTALL_ROOT>/Engine/lib/java/database*

optimSearch.jpg
Copy the file to *<DATA_EXPLORER_INSTALL_ROOT>/Engine/www/images*
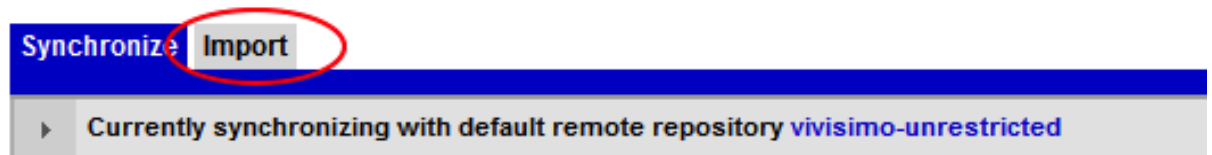
Import the Optim Search Project into Data Explorer

Once the Optim Search artifacts have been copied to the Data Explorer installation, the project itself can be imported into Data Explorer.
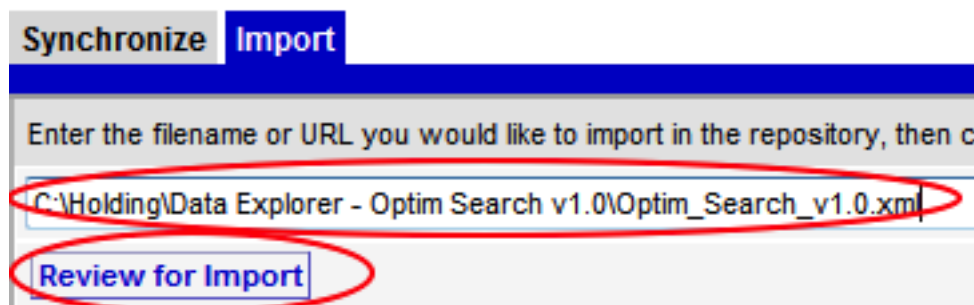
1. Logon to the Data Explorer administrative console

2. Click on the Management drop down and select Repository
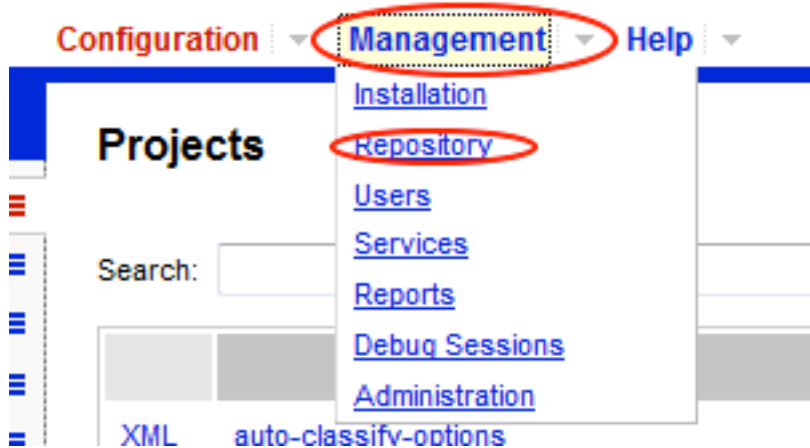
3. On the resulting screen, select the "*Import*" tab

4. In the import section, type the full path and file name to the Optim Search project XML file. (e.g.: Optim_Search_v1.0.xml)   Then select the button "*Review for Import*".

---

5. The resulting screen will show 5 nodes available for import. Select the check box next to "*Import all nodes*" and select the "*Import Selected Nodes*" button.

6. This will result in a screen confirming 5 nodes were imported. At this time, the nodes imported will no longer be shown as available for import.

Data Explorer has now been updated with an Optim Search project. See the next section to validate the Optim Search project has been imported successfully.

Validating the Optim Search Project

Once the steps above have been completed, you will have a Data Explorer project for Optim Search. To validate the project has been successfully imported complete the following steps.

1. Logon to the Data Explorer administrative console.   Ensure you are in the "*Configuration*" section of the console.

2. Select the list icon next to the Projects link in the left hand navigation pane.

3. Select the Optim Search project in the resulting page.

> **_NOTE:_** *If the "optim-search" project is not present, please review the steps in the previous section to ensure the project imported correctly.*

XML	auto-classify-options
XML	core
XML	desktop-search-security
XML	doc-search
XML	doc-search-core
XML	enron-email-tutorial
XML	iopro-preview
XML	js
XML	mobile
XML	optim-search
XML	query-meta
XML	velocity-static

4. With the Optim Search project now selected, click on the "optim-search" link in the left side of the page.

5. This will open a new window with the main Optim Search page displayed.

> ***NOTE:*** *Searching for data at this point will result in an error. This is due to the fact that the Optim Search project needs to be configured for your specific environment.*

If you were able to successfully complete all of the steps above, the Optim Search project has been installed into your Data Explorer instance. See the next section for information on how to configure Optim Search to index data for your specific installation.

CONFIGURING OPTIM SEARCH

After completing the section "*Installing Optim Search*", you are ready to modify the project to be used with your specific environment.   This section discusses how the project can be modified for your environment.

The Optim Search project imported is configured to work with the sample Optim model on a system where ODM is local to Data Explorer.   We will modify the project so it can access ODM on a remote system.   We'll examine how to change data source names, table names and SQL queries to work with your environment.

Creating a Duplicate Copy of the Optim Search Project

While you can modify the Optim Search project directly, it is recommended you make a copy of the project and modify the copy for your environment.   Creating a copy will leave your base Optim Search project exactly as imported, and allow you to reference it as needed.   Additional copies can be created as needed to meet your search needs.

> **NOTE:**   If creating a project copy from the Optim Search project is not desired, this section can be skipped.

Creating a new Collection based on the Optim Search Project

To create a copy of the Optim Search project, we will start be creating a copy of the search collection used for Optim Search.

1. Logon to the Data Explorer administration console.   Ensure you are in the "*Configuration*" section of the console.

2. On the left hand navigation panel, select the new icon for collections.

---

3. In the resulting page, enter a new Collection name and set the option "Copy defaults from" to the value "optim-odm".

> **NOTE:** It's recommended to use a naming convention that follows the pattern "optim-odm-XXXX" where XXXX is a user supplied value. (e.g.: optim-odm-copy)

New Search Collection

Create a new collection

* Name optim-odm-copy
(Name Token)

* Copy defaults from default
WAND-ontolection
ccs-collection-default
chinese-default
cs-ontolection-default
default
default-autocomplete
default-broker-push
default-push
default-tm-ontolection
example-metadata
generic-ontolection
ics-default
japanese-default
thai-default
----
auto-classify
enron-email-tutorial
iopro-sm-spotlights
iopro-tm-sample
ontolection-english-spelling-variations
ontolection-wiki-japanese-spelling
optim-odm

Description

4. When completed, click the add button.

Create a new collection   Add

* Name optim-odm-copy
(Name Token)

* Copy defaults from optim-odm

You now have a copy of the optim-odm collection.

Creating a Duplicate Copy of Optim Search

We will now create a duplicate copy of the Optim Search project. When complete, we will change the collection information to use the optim-odm collection copy (created in the steps above) as our primary search source.

1. Logon to the Data Explorer administration console. Ensure you are in the "*Configuration*" section of the console.

2. On the left hand navigation panel, select the new icon for projects.
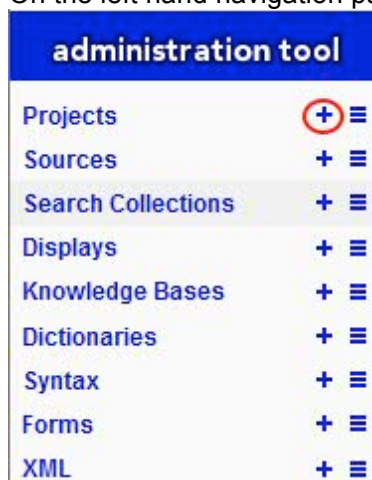
3. In the resulting page, enter a new Project name and set the option "Based on project" to the value "optim-search".

> **NOTE:** It's recommended to use a naming convention that follows the pattern "optim-search-XXXX" where XXXX is a user-supplied value.   (e.g.: optim-search-copy)

New Project

4. When completed, click the add button.

5. The new project is now completed.   However, it needs to be modified to use the new collection we created in the steps above.

6. Next to the "*Sources*" for both "*(default)*" and "*(advanced)*", select the edit button and change the name of the collection to the name you provided when creating the collection copy.   In this document, that was called "optim-odm-copy".

7. While editing, a list of valid names will be presented to ensure a valid source name is entered.   After selecting the name, click "*OK*".

8. When completed, your new project should look similar the following.

You now have a complete copy of the Optim Search project that can be configured for use in your environment.   The original project can remain without modification, and used as a reference and a base for additional project copies as needed.

Modifying the Optim Search Collection Database Seeds

> **NOTE:**   *This section assumes you are modifying the Optim Search project as imported.*
> *If you completed the previous section "**Creating a Duplicate Copy of Optim Search**", use*
> *your project and collection names in place of the default names "optim-search"*
> *"optim-odm".*

The Optim Search project contains a data source Collection called "optim-odm".   (*This is not the same as an ODM collection.   This is a Data Explorer Collection, which is a grouping of data sources to be crawled by Data Explorer.*)   In this collection there are 2 seeds for ODM that can be used for indexing Optim archived data.   The seeds are:

1. **Database (Keyed)** – A data source configuration to index all data contained within a table.
2. **Database (Custom SQL)** – A data source configuration to index data returned from a custom SQL statement.

These sources were created to provide an example of how each database seed can be used with ODM.   They can be used together, individually, or multiplied to formulate the basis of source data you wish to crawl and index.

Lets start by modifying each data source to point to an ODM instance.

Modifying the Optim Search Database Seeds
In this section, we will modify the seeds to point to an Optim ODM endpoint.

1. Logon to the Data Explorer Administrative console.   Ensure you are in the "Configuration" section of the console.

2. On the left side navigation panel, select the Search Collections list icon.
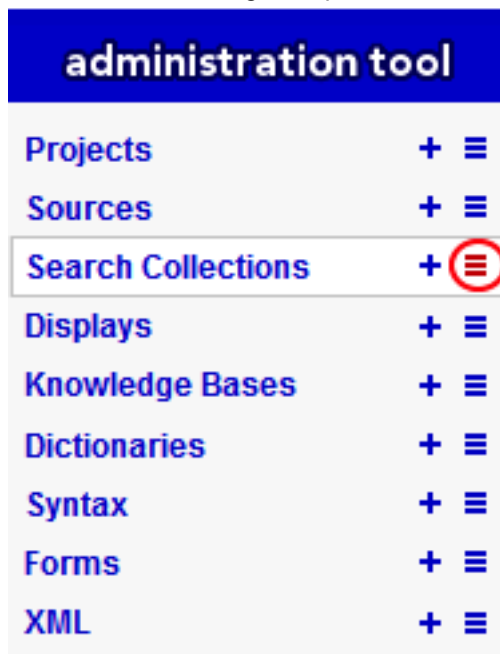
3. On the resulting page, select the "optim-odm" collection link.

> **NOTE:** Remember, if you made a copy of the collection, use the copy name you gave in place of the default name "optim-odm" listed here.

4. Click on the "Configuration" tab on the resulting screen.   This will display two database seeds. One for "*Customer SQL*", and the other for "*Keyed*".

5. To the right of the seed component named "Database (Keyed)" select the edit button.

6. This will allow you to edit the configuration.   In this example, we need to change the host name to point to an ODM endpoint that contains the data we are looking for.   (*Be sure to use a host name / IP address specific to your environment and points to where ODM is located.*)   Here, we will use an IP Address of 172.16.232.138.   Be sure to change this for both the "*Host*" and "*JDBC Connection String*" configuration parameters.

> **NOTE:**   *If a username and password are required as input in the form, and your ODM endpoint is NOT secured, use the following credentials.*
>
> *Username: anonymous*
> *Password: anonymous*

7. Next, change the "*Database name*" (ODM Data source name) and "*Table to retrieve*" parameters to match your environment.   Again, the "*JDBC Connection String*" also needs to be adjusted to include the Database name.   (e.g.:   DefTdpName=<YOUR_ODM_DATA_SOURCE_NAME>)

8. Click "Apply" and then "OK" when completed.

9. Next, repeat steps 5 thru 8 above for the Database (Custom SQL) seed.   However, in this case note the data being retrieved is based on a custom SQL statement instead of a table name. (You can change this SQL, or keep it as is if going against an archive that contains the complete Optim sample model.)

After completing the steps above your database seeds will be configured for your environment and are ready to be used to index data.

INDEXING OPTIM ARCHIVE DATA FOR SEARCH

After completing the sections above, you are ready to index and search data. In this section, we will crawl the ODM data source(s) configured in the Data Explorer Collection, and index the data. We will then use Optim Search to search the data indexed.

Crawling, Indexing & Searching Optim Archive Data

1. Logon to the Data Explorer Administrative console. Ensure you are in the "Configuration" section of the console.

2. In the left hand navigation panel, select the list icon for Search Collections.

3. In the resulting pane, select the "optim-odm" collection.

> **_NOTE:_** If you created a copy of the Optim Search project, use the collection name you specified during that process.   (e.g.   optim-odm-copy)

	Name ▲	
XML	auto-classify	
XML	ccs-collection-default	
XML	chinese-default	
XML	cs-ontolection-default	
XML	default	
XML	default-autocomplete	
XML	default-broker-push	
XML	default-push	
XML	default-tm-ontolection	
XML	enron-email-tutorial	
XML	example-metadata	
XML	generic-ontolection	
XML	ics-default	
XML	iopro-sm-spotlights	
XML	iopro-tm-sample	
XML	japanese-default	
XML	ontolection-english-spelling-variations	
XML	ontolection-wiki-japanese-spelling	
XML	optim-odm	
XML	optim-odm-copy	

4. If the resulting page indicates the Query Service is not started, click on the link provided and start the query service. When completed, select the collection again to return the collection overview. If you do not see a massage about starting the Query Service, move to the next step.

Collection optim-odm
last modified by data-explorer-admin < 1 minute ago

| Overview | Configuration | | xml | delete |

The Search Engine Query Service is not currently running. All queries on search collections and sub-collection pushes will fail until the Query Service is started.

Seeds (configuration)

odm://172.16.232.138:2551/app1/?per=1000&sql=select%20optim_customers.custname,%20optim_customers.PST_AF_NAME,%20optim_orders.order_date,%20optim_orders.order_salesman,%20optim_details.item_quantity,%20optim_items.item_description,%20optim_items.category,%20optim_items.rating,%20optim_items.unit_price%20from%20optim_customers,%20optim_orders,%20optim_details,%20optim_items%20where%20optim_customers.cust_id%20%3d%20optim_orders.cust_id%20and%20optim_orders.order_id%20%3d%20optim_details.order_id%20and%20optim_details.item_id%20%3d%20optim_items.item_id%3b&	Test it
odm://172.16.232.138:2551/app1/optim_customers?per=1000&	Test it
Live Status	start

There is no live status.

View this collection's live source

5. In the Collection Overview page, click the "*Start*" button next to "*Live Status*".

Collection optim-odm
last modified by data-explorer-admin 5 minutes ago

| Overview | Configuration | | xml | delete |

Seeds (configuration)

odm://172.16.232.138:2551/app1/?per=1000&sql=select%20optim_customers.custname,%20optim_customers.PST_AF_NAME,%20optim_orders.order_date,%20optim_orders.order_salesman,%20optim_details.item_quantity,%20optim_items.item_description,%20optim_items.category,%20optim_items.rating,%20optim_items.unit_price%20from%20optim_customers,%20optim_orders,%20optim_details,%20optim_items%20where%20optim_customers.cust_id%20%3d%20optim_orders.cust_id%20and%20optim_orders.order_id%20%3d%20optim_details.order_id%20and%20optim_details.item_id%20%3d%20optim_items.item_id%3b&	Test it
odm://172.16.232.138:2551/app1/optim_customers?per=1000&	Test it
Live Status	start

There is no live status.

View this collection's live source

6. This will start the crawling and indexing process.

7. Wen complete, the live status will look like the following.

8. In the left hand navigation panel, click on the "optim-search" link.

> **NOTE:** *If you are using a copy of Optim Search, the link will be the same name as the project name you provided.*

9. This will result in a new window with the main Optim Search page displayed. Click on the search button.
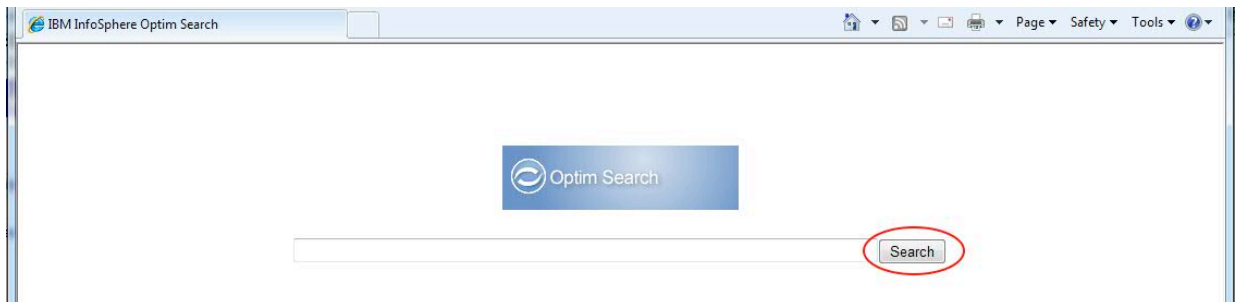
10. This will result in a seach for all items to be completed.   You will end up with a result page that looks similar to the following.   Notice that each result indicates which archive file the data resides in within your archive file population.
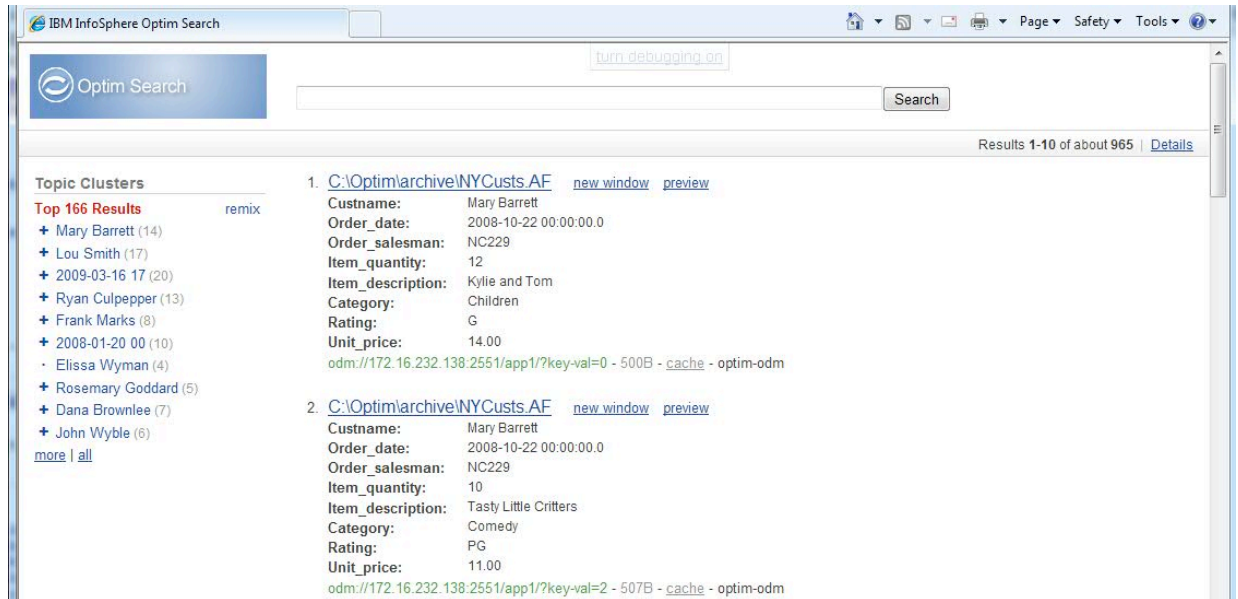
At this point, you have completed crawling, indexing and searching Optim archived data.   The Optim Search results page identifies which archive file contains the data for each result.

SUMMARY

This document has provided instructions for importing the Optim Search project into Data Explorer for use with Optim archived data.   The steps required to configure the database seeds used for ODM have been modified for your environment, and data has been crawled and indexed.   These steps have resulted in a functional Optim Search project to quickly locate data within Optim archive files.