TRIAL ACCESS GUIDE

infoarchive by EMC
# Table of Contents

- Overview .................................................................................................................................................. 3
- InfoArchive User Interface Overview ....................................................................................................... 3
- ETL / Data Import Example ...................................................................................................................... 9
- Web UI Modification ................................................................................................................................. 20
- Reporting via Third-Party Tools ................................................................................................................ 25
Overview

Welcome to the EMC InfoArchive 3.2 Cloud Trial Walkthrough. The steps in this document will familiarize you with the application decommissioning capabilities of EMC InfoArchive.

InfoArchive User Interface Overview

In this scenario, you’ll experience the web-based application used by business users to access information which is stored in the InfoArchive repository. This is a lightweight user interface which can be tailored (via configuration) to the needs of the business. This configuration is done via established standards such as HTML, XML, CSS, etc.

1. From the Windows desktop, double-click the shortcut for EMC InfoArchive

2. Google Chrome will launch and you will be presented with the InfoArchive login screen

3. Provide a username of admin and a password of secret to sign in

4. Select the Employee Management node beneath InfoArchive to search against employee data stored in the archive
The nodes displayed here represent data which has been archived from different source applications.

5. Click the **Submit** button with no criteria to return a list of all employee data stored in the archive.
6. Click the value for first name (such as “JASON”) to select a particular employee returned by the search.

![Search Result Table]

The list of columns, nested searches, and search fields are configurable as we will see later in the tutorial.

7. More detailed information is displayed in the adjacent Record Details pane.

![Record Details Pane]
8. Click the Expand button to maximize the area used to display the selected record details

The layout and style of the data is configurable using XSL transformation stylesheets.

9. Click the X icon to hide the expanded view
10. Click the Export PDF button to render the detailed employee information to a downloadable PDF.

11. Click the downloaded PDF and it will display in the browser. Close the tab when finished.

12. In the Search Result pane, click the icon to display archived Pay History information associated with a particular employee.

The icon provides an easy way to execute a separate, but related search without needing to type or otherwise copy-and-paste information.
13. All pay history records for the employee are displayed in the Search Result pane.

14. Select a specific pay history record.

15. As before, more detailed information for the record is displayed in the **Record Details** pane.

16. Leave the browser open for the next section.

This concludes the UI overview section.
ETL / Data Import Example

In this section, you'll perform an “extract, transform and load” (ETL) of data. Data will be extracted from a source system, namely Microsoft SQL Server. The data transformation will entail the conversion of relational database rows into an XML document. And finally the data will be loaded into the target system, namely EMC InfoArchive.

The software tool used in this guide to perform the ETL is Talend Open Studio, which is a commercial, open source data integration product developed by Talend (http://www.talend.com/). However, it is important to note that there are many ETL tools available on the market, any of which can be used in conjunction with EMC InfoArchive. In addition to pre-built ETL tools, custom solutions can be utilized to get data into InfoArchive.

1. As you left the browser open from the previous session, click the Employee Search tab in the Search pane.

2. In the Last Name search field, type Burton and click the Submit button.
3. Notice that no search results are returned

The employee archive currently contains no references to any employees with the last name “Burton.” However, you will be performing an import of additional data into the archive. And at the end of this section, you will perform this same search a second time with different results.

4. Use the shortcut on the desktop to open Talend Open Studio

5. Ensure that the talendxdb project is selected and click Open

6. If prompted to connect to the Talend online community, press Skip
7. Once Talend Open Studio loads, expand the **Job Designs** node in the **Repository** pane.

8. Double-click the **Employees_SQLtoInfoArchive_File 0.1** job design to open it for execution.

   ![Diagram of Talend Open Studio interface showing the Business Models and Job Designs nodes]

   This job design has been configured to pull data from a specific Microsoft SQL Server repository, using a specific SQL query. It then transforms the data into a XML format using a temporary file, and then loads it into the specified InfoArchive instance. As mentioned above, virtually any ETL tool can be used to perform these actions.
9. We will now walk through the configuration of the job design prior to execution. Single-click to select the tMSSqlInput_2 activity.

10. In the Component pane below the job design view, you can see the extraction details. Including the connection info of the Microsoft SQL Server, and the SQL query being used (if you scroll down).
11. Double-click the `tAdvancedFileOutput_XML_1` activity to display the transformation details.

12. On the left, we see the columns being returned from the SQL database. On the right, the desired XML structure to be used for each row returned by the query. Click **Ok** to close the dialog when you’re done examining.
13. Single-click the txDBPut_1 activity to display the information used to load the XML into InfoArchive.

14. In the Component pane below the job design view, see the connection details for the InfoArchive xDB repository.

Note the Path configuration. This is the location within the InfoArchive repository where the XML data will be stored.
15. Now we’ll actually execute the job design. On the toolbar, click the Run icon

16. The execution progress shows below in the Run pane

The job is complete when you see the exit code of 0, as above.
17. Now we will verify that the data was successfully imported using the InfoArchive xDB Admin Client. Use the following icon on the Windows task bar to open the client.

![Windows task bar icon](image)

18. From the **Settings** menu, click **Select active federation**

![Select active federation](image)

19. On the Select federation dialog, add the federation name by appending `#infoarchive_ad` to the end of the bootstrap path as seen below. Then click **OK**.

![Select federation](image)

A federation is a container for related databases, to which it provides a single server connection. A federation can contain as many databases as needed.
20. Click the connect icon to bring up the connection dialog

21. Enter a password of **secret** and click **OK** to connect
22. Expand the necessary nodes to navigate to root-library / DATA / infoarchive_ad / Collection / PAY_HISHDR / employee_batch_1.xml

This is the XML data which was added by Talend Open Studio. Above, PAY_HISHDR.XML represents the pre-existing employee data. Both data sets are appended when accessing the employee data from the InfoArchive web user interface.
23. Return to the web browser, and re-run the Employee Search with a last name of Burton

![Employee Search](image)

24. The new record is returned in the search result pane

![Search Result](image)

This concludes the ETL / data import example.
Web UI Modification

In this section you will be making a modification to the InfoArchive user interface. Specifically, you will be extending the Employee Search to allow searching against the First Name field in the repository. The tool you will be using to perform this modification is Eclipse. Though, any text or XML editor can be used. In a future release of InfoArchive, these changes will be possible via a graphical configuration tool.

1. Open Eclipse using the icon on the Windows desktop

2. In the Package Explorer pane, navigate to DDS / applications / infoarchive_ad / data / application / resources / template

3. Double-click template-content.xml for editing
4. From the **Edit** menu, select **Find/Replace**

5. Type **Remove Comments** in the Find field and click the **Find** button
6. Delete the XML comment markers before and after the opening and closing searchField tags. The end result should look like the following:

```xml
<whereclauseparameter>$elem/LAST_NAME$</whereclauseparameter>
<fullText>true</fullText>
</searchField>

<!--Remove Comments to add search field -->

<searchField>
  <name>first_name</name>
  <label>First Name</label>
  <whereclauseparameter>$elem/FIRST_NAME$</whereclauseparameter>
  <fullText>true</fullText>
</searchField>

<searchField>
  <name>position</name>
  <label>Position</label>
</searchField>
```

7. **Save** the file by using Ctrl-S or using the save icon on the toolbar.

8. In the Runner pane, expand the InfoArchive node, and double-click the contained run configuration.

   ![Runner Pane]

   This run configuration will execute the necessary script to bundle and deploy any changes which were made to the InfoArchive UI application.
9. Once the run configuration is done, a browser window will open targeting the InfoArchive web UI. Sign as admin with a password of secret.

10. In the InfoArchive App Decom pane, select the Employee Management node.
11. In the **Search** Pane, select the **Employee Search** tab, type **Tim** for the value of **First Name** and click **Submit**

![Search Pane with Employee Search tab selected](image)

The “First Name” search field was previously unavailable, prior to performing the steps in the exercise.

12. The relevant result is displayed in the **Search Result** pane

![Search Result pane](image)

This concludes the Web UI Modification scenario.
Reporting via Third-Party Tools

In this scenario you will see an example of how third-party reporting tools can be used. Selected portions of the InfoArchive repository can be exposed to JDBC-capable tools, such as Crystal Reports, and TIBCO Jaspersoft Studio (http://www.jaspersoft.com/).

1. Use the icon on the Windows task bar to open Jaspersoft Studio

2. In the Repository pane, double-click InfoArchive JDBC data adapter

3. Note the connection details. A special JDBC driver is provided for InfoArchive xDB connectivity which allows interaction with selected areas of the repository
4. Select the Driver Classpath tab to view the file location of the InfoArchive JDBC driver and its associated library. These JARs can be added to any JDBC-enabled reporting tools.

5. Click **Cancel** to close the Data Adapter Wizard dialog.
6. With the Employees.jrxml report open, select the Source tab at the bottom, then note the SQL query being used for the report.

```xml
<jasperReport xmlns="http://jasperreports.sourceforge.net/jrxml">
  <queryString>
    <![CDATA[select * from EMPLOYEES.EMPLOYEES]]>
  </queryString>
  <field name="emp_no" class="java.lang.String"/>
  <field name="first_name" class="java.lang.String"/>
  <field name="last_name" class="java.lang.String"/>
  <field name="gender" class="java.lang.String"/>
  <field name="birth_date" class="java.sql.Date"/>
  <field name="hire_date" class="java.sql.Date"/>
</jasperReport>
```

In this scenario, we’re simply grabbing all of the data in the employees dataset from InfoArchive. More complex queries, including JOIN and WHERE clauses, can be utilized as necessary.
7. Click the Preview tab at the bottom to execute the query and display the report after some time

8. The first page is displayed as additional pages continue to be generated

<table>
<thead>
<tr>
<th>Emp #</th>
<th>First Name</th>
<th>Last Name</th>
<th>Gender</th>
<th>Birth Date</th>
<th>Hire Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>10001</td>
<td>Georgi</td>
<td>Facello</td>
<td>M</td>
<td>Sep 02, 1953</td>
<td>Jun 26, 1986</td>
</tr>
<tr>
<td>10002</td>
<td>Bezalol</td>
<td>Simmel</td>
<td>F</td>
<td>Jun 02, 1964</td>
<td>Nov 21, 1985</td>
</tr>
<tr>
<td>10003</td>
<td>Parto</td>
<td>Barnford</td>
<td>M</td>
<td>Dec 03, 1959</td>
<td>Aug 28, 1986</td>
</tr>
<tr>
<td>10004</td>
<td>Chirstian</td>
<td>Koblick</td>
<td>M</td>
<td>May 01, 1954</td>
<td>Dec 01, 1986</td>
</tr>
<tr>
<td>10005</td>
<td>Kyoichi</td>
<td>Maliniak</td>
<td>M</td>
<td>Jan 21, 1955</td>
<td>Sep 12, 1989</td>
</tr>
<tr>
<td>10006</td>
<td>Anneke</td>
<td>Preusig</td>
<td>F</td>
<td>Apr 20, 1953</td>
<td>Jun 02, 1989</td>
</tr>
<tr>
<td>10007</td>
<td>Tzvetan</td>
<td>Zielinski</td>
<td>F</td>
<td>May 23, 1957</td>
<td>Feb 10, 1989</td>
</tr>
<tr>
<td>10008</td>
<td>Saniya</td>
<td>Kalloufi</td>
<td>M</td>
<td>Feb 19, 1958</td>
<td>Sep 15, 1994</td>
</tr>
<tr>
<td>10009</td>
<td>Sumant</td>
<td>Peac</td>
<td>F</td>
<td>Apr 19, 1952</td>
<td>Feb 18, 1985</td>
</tr>
<tr>
<td>10010</td>
<td>Duangkaew</td>
<td>Piveteau</td>
<td>F</td>
<td>Jun 01, 1963</td>
<td>Aug 24, 1989</td>
</tr>
<tr>
<td>10011</td>
<td>Mary</td>
<td>Sluis</td>
<td>F</td>
<td>Nov 07, 1953</td>
<td>Jan 22, 1990</td>
</tr>
<tr>
<td>10012</td>
<td>Patricio</td>
<td>Bridgland</td>
<td>M</td>
<td>Oct 04, 1960</td>
<td>Dec 18, 1992</td>
</tr>
</tbody>
</table>

9. Press the Stop icon to cancel execution of the report
10. An additional report, Salaries.jrxml is also provided for your review. Repeat steps 6-9 for this report if desired. This concludes the EMC InfoArchive Cloud Trial Walkthrough.