Informatica

Informatica® Data Archive
6.4.3

Retiring an Application to Data Vault With On-Premise Data Archive

Informatica Data Archive Retiring an Application to Data Vault With On-Premise Data Archive
6.4.3

Publication Date: 2018-06-09

Table of Contents

Abstract

You can use on-premise Data Archive to retire an application to the Informatica Data Vault. This article describes the steps to import metadata, create entities, and create and schedule a retirement project.

CHAPTER 1

Retire an Application to Data Vault With On-Premise Data Archive

This chapter includes the following topics:

Application Retirement Overview

Use on-premise Data Archive to retire a legacy application to Data Vault. After you retire the application, you can continue to access the retired data and enable compliance.

When you retire an application to Data Vault, you copy the transaction tables, master tables, and metadata from a structured application on your source database to a compressed file format in Data Vault.

To retire data, you first import the metadata. Then, you create retirement entities. Each entity represents a collection of tables. You create a retirement project where you specify the entities to retire. When you schedule the retirement project, the Data Vault Loader job loads the entities to Data Vault.

After you retire the application to Data Vault, you can access and manage the retired data.

This article explains the retirement process with an example.

Prerequisites

Before you begin the retirement process, verify the following prerequisites:

1. Create database users and assign privileges.
2. Configure properties in the `conf.properties` file.
3. Start the ILM application server.
4. Configure system profile properties on the **Configuration Settings** page in Data Archive.

For more information about the prerequisite tasks, see the *Informatica Data Archive Administrator Guide*.

Application Retirement Process

To retire an application, you perform tasks through the Enterprise Data Manager and Data Archive user interface.

1. Start Enterprise Data Manager.
2. Create an application version.
3. Import the metadata from the source database.
4. Create the retirement entities.
5. Create a source connection in Data Archive.
6. Create a target connection in Data Archive.
7. Create a Data Vault access role.
8. Assign the Data Vault access role to the user.
9. Create and schedule a retirement project.
10. Review the row count report.
11. Validate data through the Data Discovery portal.

Airline Ticket Example

This article uses an example to explain the retirement process.

You use an application called Airline for airline ticket details. The application contains 215 tables and resides on an Oracle source database. You want to start using a different application for airline ticket information. You plan to retire the Airline application to Data Vault for the following reasons:

- Retain the data for compliance reasons.
- Improve performance on the source database.
- Eliminate cost to support the Airline application.
- Maintain access to the Airline application.

After you use Data Archive to retire the application, the Airline application still exists on the source database. You plan to decommission the application in accordance with your company policy.

The following table lists a few of the tables and columns in the application that you plan to retire:

Table	Columns
TICKET_NUMBER	- TICKET_NUMBER - ISSUE_DATE - CUST_ID
TICKET_COUPON	- TICKET_NUBMER - COUPON - FROM_CITY - TO_CITY
TICKET_COUPON_SAVE	- TICKET_NUMBER - COUPON - IS_USED
TICKET_PARTITION	- TICKET_NUMBER - ISSUE_DATE - PASSENGER_NAME
TICKET_PARTITION_TARGET	- TICKET_NUMBER - PASSENGER_NAME - ID_NUM
TICKET_SAVE	- TICKET_NUMBER - ISSUE_DATE - CREDITCARD

Step 1. Start Enterprise Data Manager

Open Enterprise Data Manager from Data Archive.

1. In Data Archive, select **Accelerators** > **Enterprise Data Manager**.

   The **Save As** window opens.
2. Save `RunEdm.jnlp` on the local machine.

3. To open Enterprise Data Manager, run `RunEdm.jnlp`.

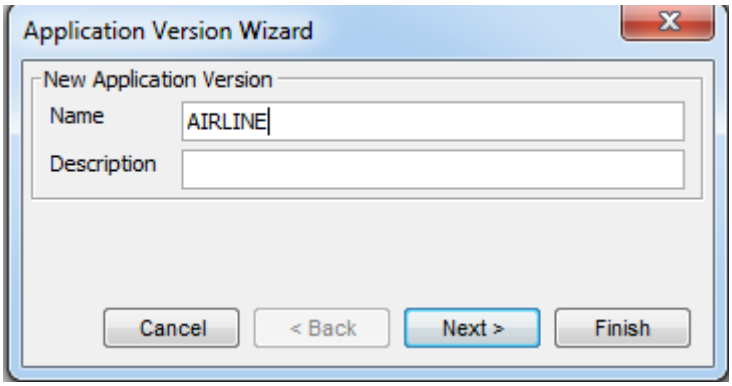
   Note: You must have Java 7 installed on the machine where you run Enterprise Data Manager.

Step 2. Create an Application Version

Create an application version for the legacy application that you plan to retire.

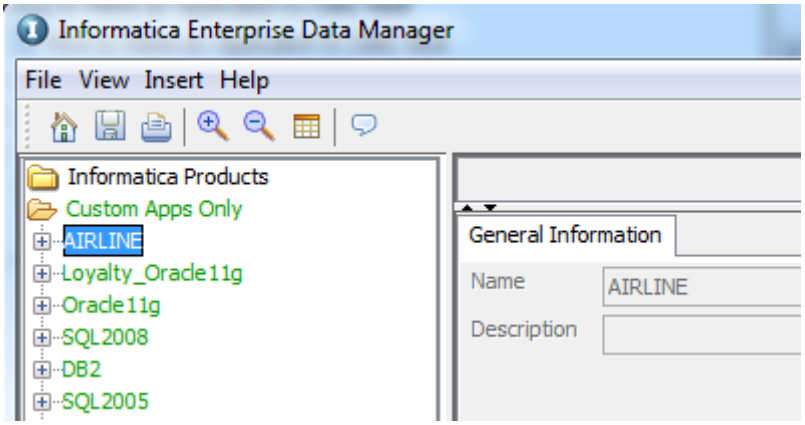
1. In Enterprise Data Manager, right-click **Custom Apps Only** on the **Explorer** pane.
2. Select **New Application Version**.
3. Enter the application version name.

   In this example, the application version name is AIRLINE.



4. Click **Finish**.
5. Click the **Save** icon.

   The AIRLINE application version appears below the Custom Apps Only application.



Step 3. Import Metadata from the Source Database

To import table metadata, first define a connection to the source database and then specify the application module to import.

When you import metadata, you import a list of tables and the table constraints from the source database.

1. In Enterprise Data Manager, expand the application **Custom Apps Only** on the **Explorer** pane.

   The list of application versions appear.
2. Click the application version that you created.
3. Click **File** > **Import Metadata from Database**.

   The **Connect to Import Metadata from Database** window appears.
4. In the **Connect to Import Metadata from Database** window, you can use the **Connection Name** parameter to select a connection that you previously defined, or you can create a connection. Configure the following parameters to create a connection:

   Database Type

   The type of database that you want to connect to.

   Database Host

   The name of the machine that hosts the source data.

   Database Port

   The port number that the source database uses.

   Service Name

   Unique identifier or system identifier for the source database server.

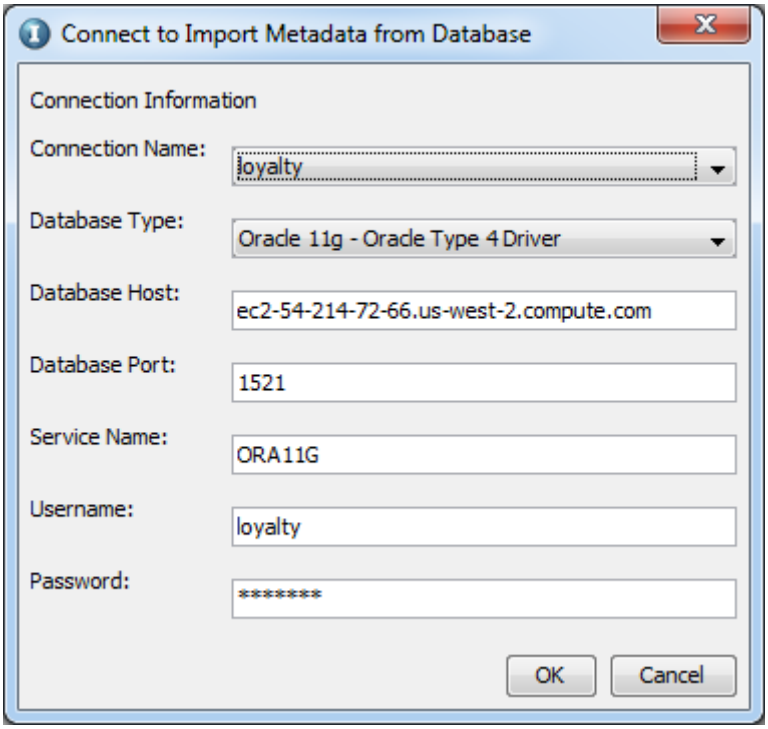
   User Name

   The name of the user that connects to the source database. The user is the owner of the application module and tables that you want to import the metadata for.
   For Oracle, the user name is the same as the name of the application module.

   Password

   The password for the database user name.

In the airline example, you name the connection loyalty and enter details to the source database.

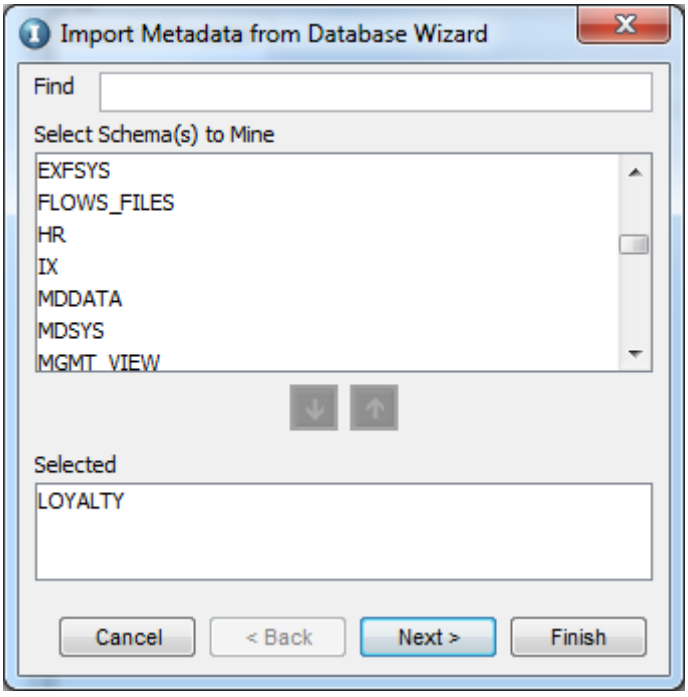


5. Click **OK**.

   The **Import Metadata from Database Wizard** opens.

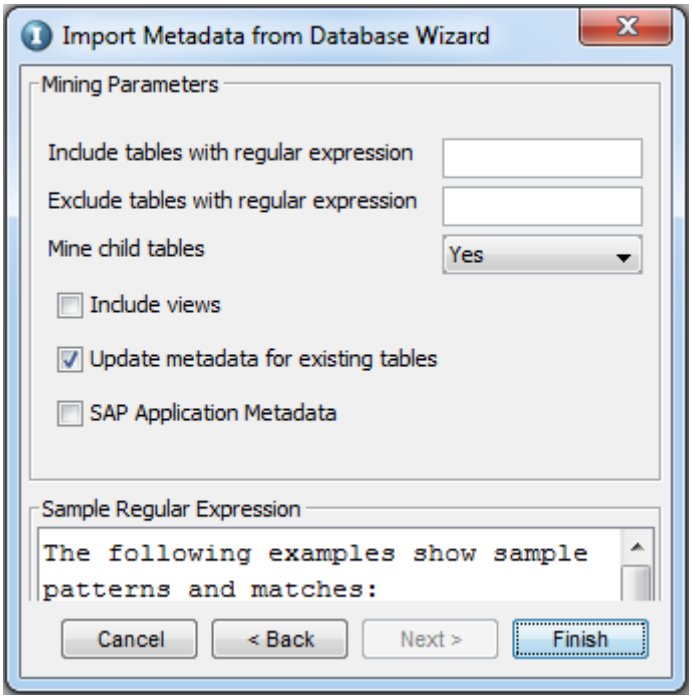
6. Select the schemas that you want to import metadata from. Double-click a schema to move it to the **Selected** box. Alternatively, select the schema and use the down arrow.

   In the airline example, the legacy application contains one schema called LOYALTY.



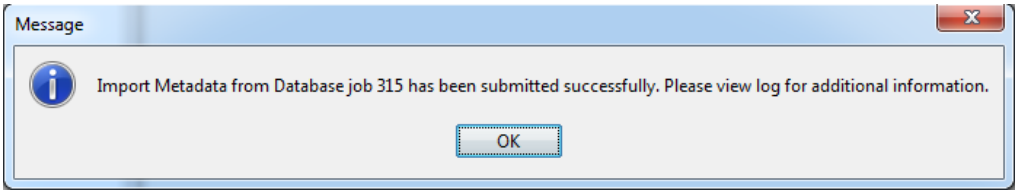
7. Click **Next**.
8. Choose one of the following options:
   - **Submit Import Metadata as a Background Job**.
     Runs the job in the background. If you run the job in the background, you can continue to perform other tasks.
   - **Continue Import Metadata through EDM**.
     Runs the job in the foreground. If you run the job in the foreground, you must wait until the job completes to perform another task. Additionally, if you run the job in the foreground and you have a large volume of metadata to import, you might receive memory errors. You might want to run the job in the foreground if you have a low volume of metadata to import.

   Typically, an application marked for retirement contains a large number of tables and a large volume of metadata. For this reason, the recommended best practice is to import metadata as a background job. For the airline example, choose **Submit Import Metadata as a Background Job**.
9. Click **Next**.
10. Select **Update metadata for existing tables**.



11. Click **Finish**.

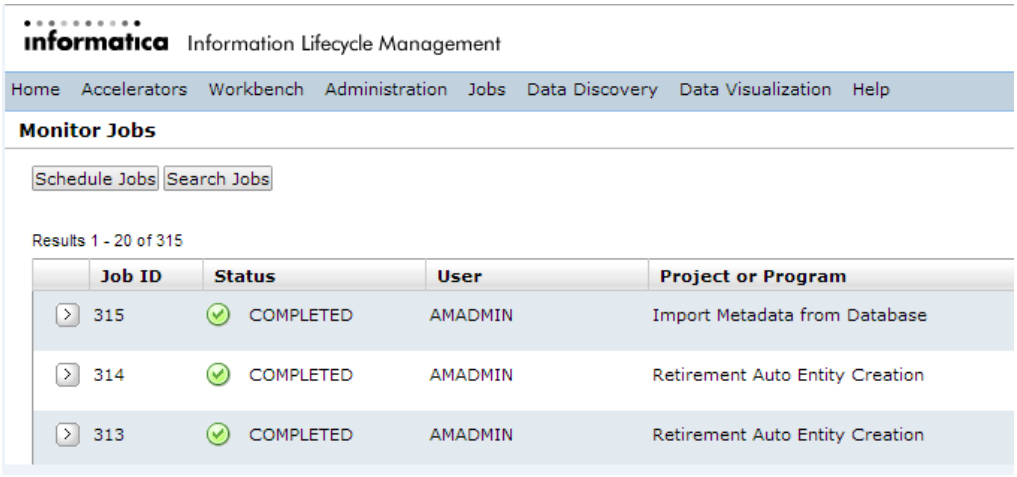
    You receive a message that the wizard started the Import Metadata from Database job. You can use the job ID in the message to monitor the job status in Data Archive.



12. Click **OK** to close the message window.
13. To view the status of the job, go to the Data Archive user interface.

14. Click **Jobs** > **Monitor Jobs**.

    When the job is complete, the job status is COMPLETED. You can expand the job icon on the left to view job details.



Step 4. Create Retirement Entities

You must create one or more retirement entities to retire the tables in the legacy application. The recommended number of tables in an entity is 200. If you have 200 tables or less in the legacy application, you can create one retirement entity with all the tables. If you have more than 200 tables, use the **Generate Retirement Entity** wizard to generate retirement entities.

You create an entity to represent a group of tables. A retirement entity contains a group of tables that might or might not be related. The entity can contain tables from multiple schemas. Retirement entities do not require constraint information.

In the airline example, the LOYALTY application module contains 215 tables. The following section describes how to generate multiple retirement entities from the tables in the LOYALTY application module.

Generating Retirement Entities

The **Generate Retirement Entity** wizard runs the Retirement Auto Entity Creation job. You specify the number of tables that you want in each entity. You can also choose to add tables from different schemas in the same entity.

1. In Enterprise Data Manager, click **View** > **Database Archive**.
2. Navigate to the application version that you created in Step 2.
3. Right-click on the application version and choose **Generate Retirement Entities**.

   The **Generate Retirement Entity** wizard appears.
4. Enter the following retirement entity generation parameters:

   Number of Tables in each Entity

   Required. Maximum number of tables to include in each entity.

Prefix for Entity

Required. Text to append before the entity name. You can use the application name as the prefix. When you create the retirement project, you can look for entities that start with the same prefix and add them to the project.
The full entity name, including the prefix and suffix, cannot exceed 45 characters.

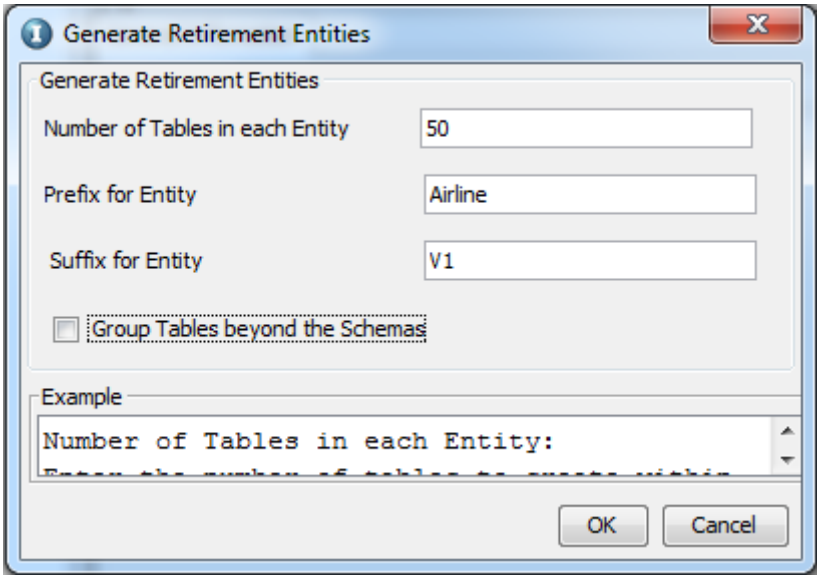
Suffix for Entity

Optional. Text to append after the entity name. The full entity name, including the prefix and suffix, cannot exceed 45 characters.

Group Tables beyond the Schemas

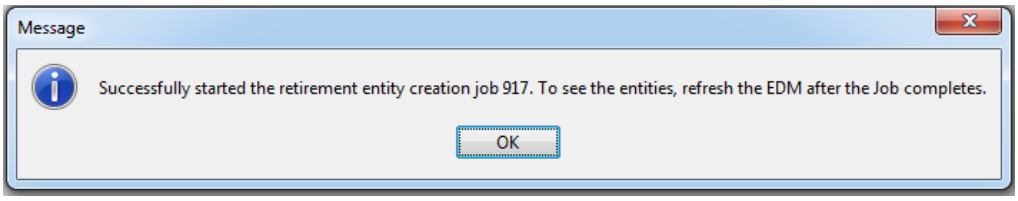
Optional. Determines whether the entity can include tables from multiple schemas. If enabled, the wizard creates the maximum number of tables that you specify from across schemas. If disabled, the wizard creates the entity with the maximum number of tables within one schema. Default is enabled.

In the Airline example, the application contains 215 tables. The following image shows the entity specification. Each entity will contain a maximum of 50 tables. The prefix for each entity name is Airline to match the application version name. The suffix for each entity name is V1 to indicate the version number of the application.



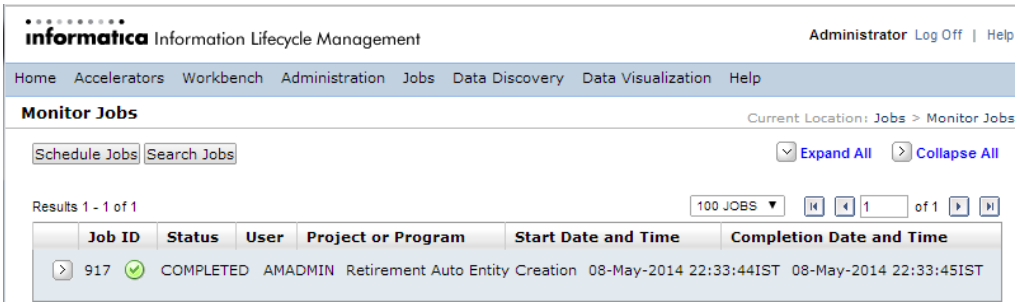
5. Click **OK**.

   You receive a message that the wizard started the Retirement Entity Creation job. You can use the job ID in the message to monitor the job status in Data Archive.



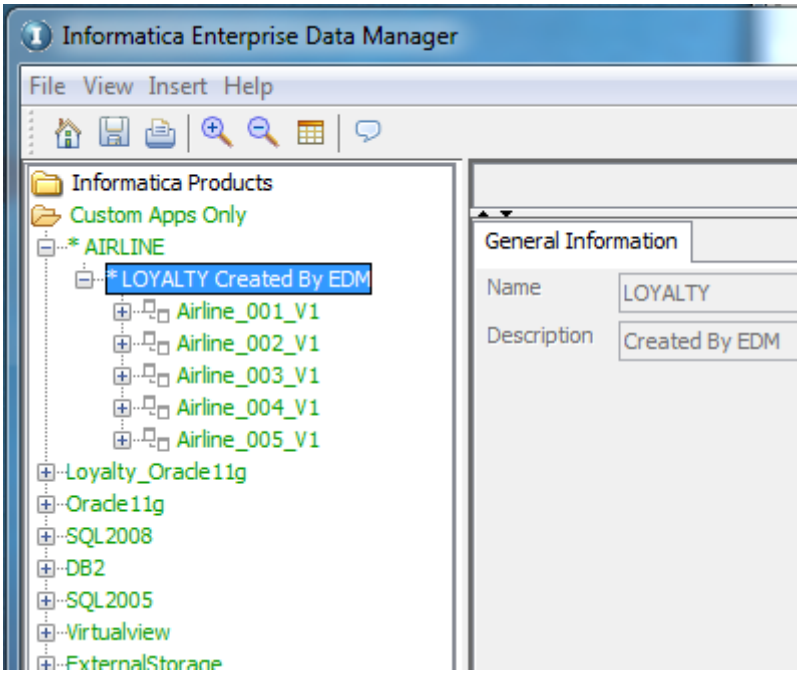
6. Click **OK** to close the message window.
7. To view the status of the job, go to the Data Archive user interface.
8. Click **Jobs** > **Monitor Jobs**.

When the job is complete, the job status is COMPLETED. You can expand the job icon on the left to view job details such as the list of entities the job created.



9. In Enterprise Data Manager, click **View** > **Database Archive**.
10. Navigate to the application module.
11. Expand the application module.

    The list of retirement entities appear in the Loyalty module.



Step 5. Create a Source Connection

Create a connection to the source database that stores the application that you want to retire.

When you create a source connection, you define general properties and database properties. The Data Vault Loader job uses the attributes that you specify to connect to the source database. Use the same source connection information you entered to import the metadata in Enterprise Data Manager.

1. Go to the Data Archive user interface.
2. Choose one of the following options:

- Click **Administration** > **New Source Connection**.
- If you have an existing source connection, click **Administrator** > **Manage Connections**. Click the **Source** tab. Click the copy icon for the source connection you want to copy.

The **Create or Edit an Archive Source** page appears.

3. Define the following general properties:

Connection Name

Name of the source connection. When you create a retirement project and you assign a source, you select the connection name that you define here.

Connection Description

Long text description for the connection name. When you manage connections, you can use this field to filter the list of connections.

Connection Type

Database connection type that determines how you connect to the source database.

Application Version

Business application and application versions that the database connection type supports. The possible values depend on the database connection type.

4. Define the following database properties:

Host

IP address of the source application database server. When the source application server is in an Oracle Real Application Cluster, enter the IP address of node1 and the IP address of node 2, separated by a comma.

Port

Port of the source application database server. When the source application server is in an Oracle RAC, enter the port number of node 1 and port number of node 2, separated by a comma.

Service Name

Unique identifier or system identifier for the source application database server.

When you define an Oracle RAC source database server, enter the service name.

Admin Schema Name

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Admin Login Name is the same as the schema name.

Password

Password for the administration login name.

Apps Schema Name

Database user that owns the tables in the application that you want retire. For example, APPS for Oracle applications.

Application Login Name

Login name that connects to the source database that contains the data that you want to retire. This user does not require special permissions as this value is used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Staging Schema Name

Staging database user for the source database. However, the retirement process does not use a staging database.
Enter the value you entered for the **Apps Schema Name**.

Staging Login Name

Login name for the staging database. However, the retirement process does not use a staging database.
Enter the value you entered for the **Application Login Name**.

Password

Password for the staging login name. However, the retirement process does not use a staging database.
Enter the value you entered for the application login **Password**.

Staging Tablespace

Tablespace in the staging database. Enter the name of the tablespace you created when you installed Data Archive. For example, AM_DATA.

Use Copy to Staging

Determines whether an archive job temporarily stores data in a staging area before it moves data to the target.

Clear the check box. Data Archive does not store data in a staging area during the retirement process.

Use Row ID for Delete

Determines how the archive job removes data from the source database.

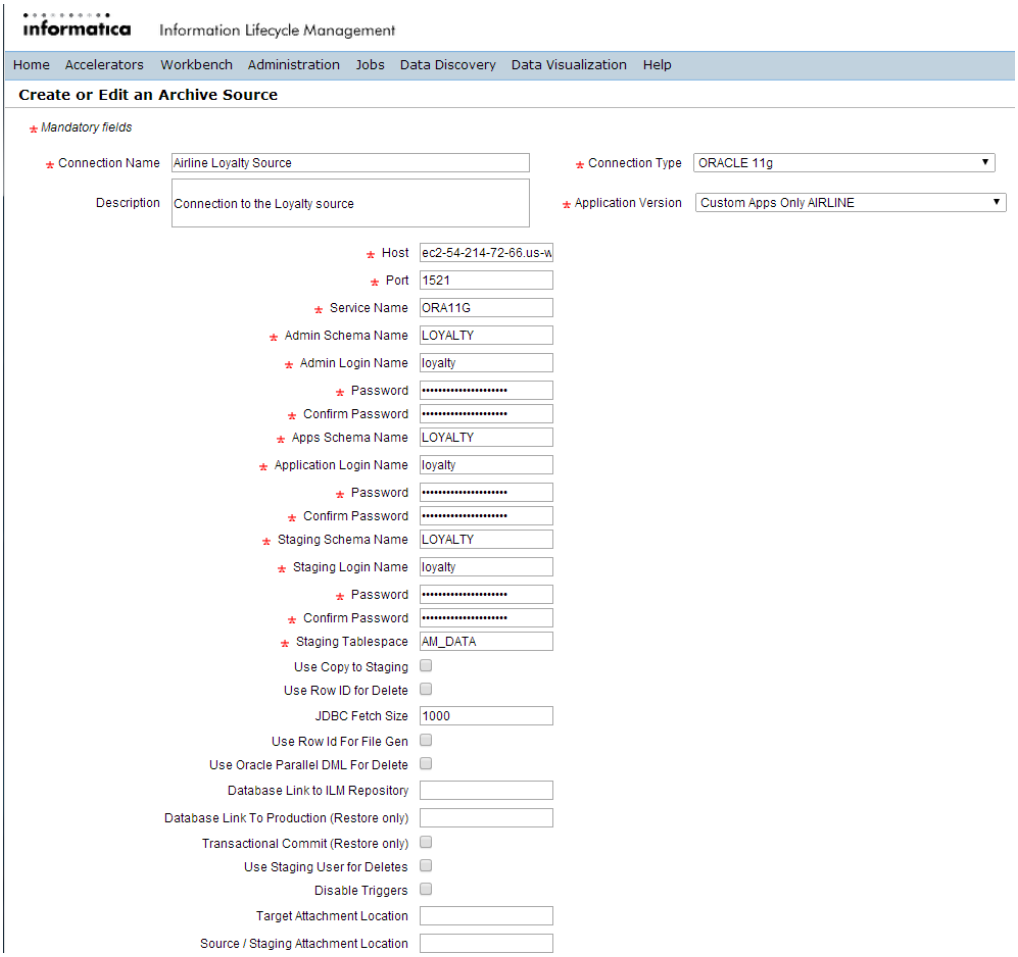
Clear the check box. Data Archive does not remove data from the source database during the retirement process.

Use Row ID for File Gen

Determines how the archive job generates the file from the source database. When you enable this property, multiple select statements act on the same table causing a number of database connections. Multiple database connections might impact the performance of the source database. For application retirement, the recommended best practice is to use one select statement on each source table.

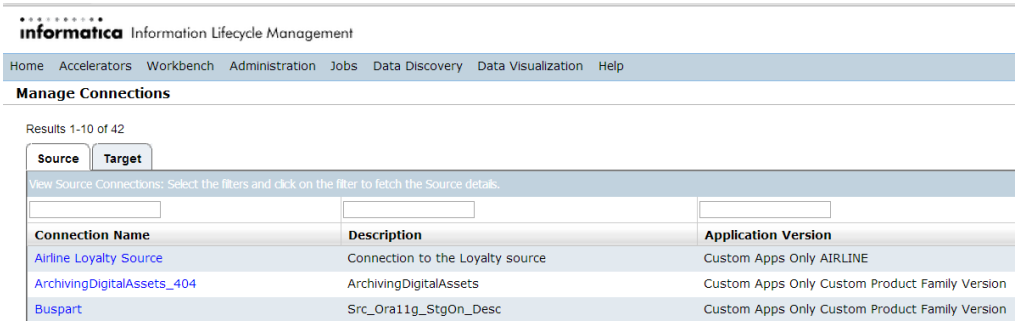
Clear the check box.

Note: You do not need to enter or update values for the other parameters on the **Create or Edit an Archive Source** page.



5. Click **Save**.

   The system validates the database connection. The **Manage Connections** page appears with the source connection listed on the **Source** tab.



Step 6. Create a Target Connection

Create a connection to the Data Vault, the target destination for the legacy application.

When you create a target connection, you define general properties and database properties. The Data Vault Loader job uses the attributes that you specify to connect to the target.

1. Go to the Data Archive user interface.
2. Choose one of the following options:
   - Click **Administration** > **New Target Connection**.
   - If you have an existing target connection, click **Administrator** > **Manage Connections**. Click the **Target** tab. Click the copy icon for the target connection you want to copy.

   The **Create or Edit an Archive Target** page appears.
3. Define the following general properties:

   Connection Name

   Name of the target connection. When you create a retirement project and you assign a target, you select the connection name that you define here.

   Connection Description

   Long text description for the connection name. When you manage connections, you can use this field to filter the list of connections.

   Connection Type

   Database connection type that determines how you connect to the target database.

   Application Version

   Business application and application versions that the database connection type supports. The possible values depend on the database connection type.
4. Define the following database properties:

   Staging Directory

   Directory in which the Data Vault Loader job temporarily stores data as it completes the retirement process. Enter the absolute path for the directory.
   The directory must be accessible to the ILM application server.

   Number of Rows Per File

   Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is 1 million rows.

   Data Vault Data Directory

   Directory in which the Data Vault Loader job creates the archive folder. Enter the absolute path for the directory.

   Data Vault Archive Folder Name

   Name of the archive folder in Data Vault where the application data is stored. The Data Vault archive folder corresponds to the database in the source.

   Data Vault Host

   Host name or IP address of the machine that hosts the Data Vault Service.

Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

Data Vault User

Name of the administrator user account to connect to the Data Vault Service.
You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

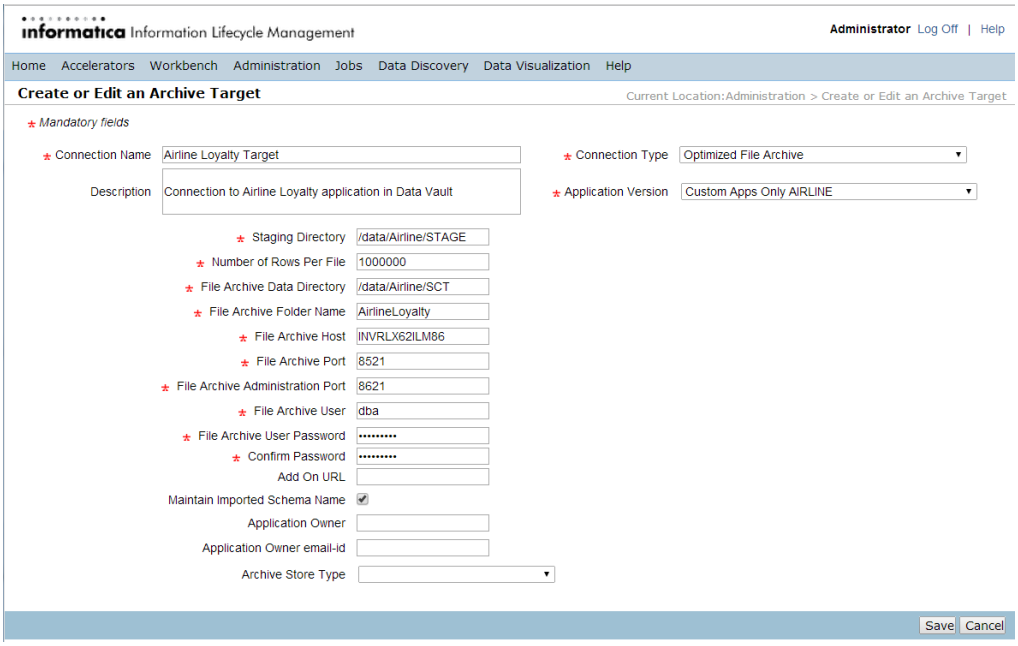
Data Vault User Password

Password for the administrator user account.

Confirm Password

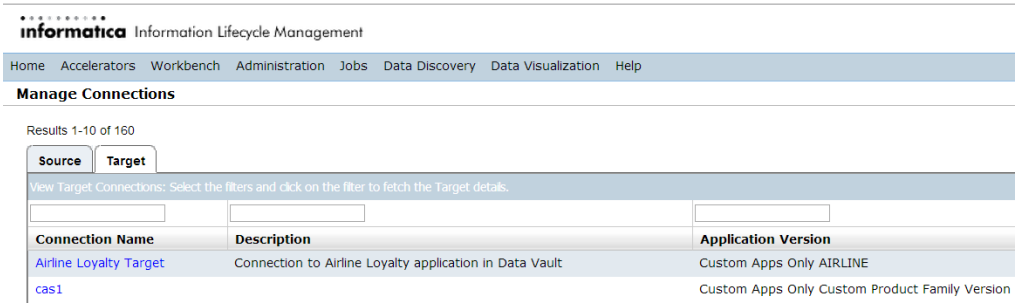
Verification of the password for the administrator user account.

Note: You do not need to enter or update values for the other parameters on the **Create or Edit an Archive Target** page.



5. Click **Save**.

   The system validates the database connection. The **Manage Connections** page appears with the target connection listed on the **Target** tab.



Step 7. Create a Data Vault Access Role

To validate data through the Data Discovery portal, you create a Data Vault access role. After you create the Data Vault access role, you assign the role to the user who performs the data validation step. You also assign the role to each entity when you create the retirement project. Users with an access role can view archived data in entities with the same access role.

When you create a Data Vault access role, you define the name, description, and a period of time that the role is valid.

1. Click **Administration** > **Manage Roles**.
2. Click **New Access Role**.

   The role name cannot contain special characters.

   The **Create or Edit a Role** page appears.
3. Enter the following Data Vault access role properties:

   Role Name

   Name of the role.

   Description

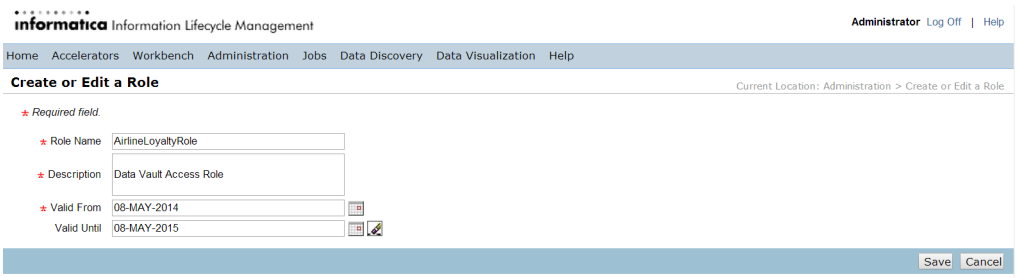
   Description of the role. When you assign the role to a user, the description appears in the drop-down menu of available roles.

   Valid from

   Start date of the user account validity period. Click the calendar icon to select a date.

Valid until

End date of the user account validity period. Click the calendar icon to select a date.



4. Click **Save**.

   The role appears on the **Manage Roles** page.

After you create the access role, assign the role to a user and the entities that you want to retire.

Step 8. Assign the Data Vault Access Role to User

After you create the Data Vault access role, assign the role to a user. You can assign the role to the default administrator user, AMADMIN, or to a user that you create for retirement.

1. Click **Administration** > **Manage Users**.

   The **Manage Users** page appears.

2. Click the **Edit** icon next to the user that you want to assign the access role to.

   The **Edit Users** page appears.

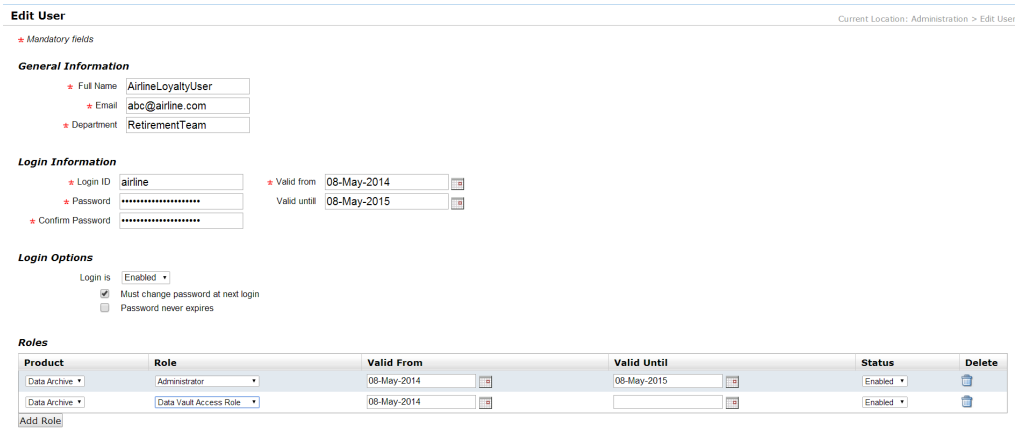
3. Click **Add Role.**

   A new line appears in the **Roles** table.

4. From the **Role** menu, select the role that you want to assign to the user.

   Note: The roles appear in the menu as the description you entered when you created the role.

   In the airline example, you assign the Data Vault access role AirlineLoyaltyRole, to the user AirlineLoyaltyUser.



5. Click **Save**.

Step 9. Create and Schedule a Retirement Project

You create a retirement project to specify the application to retire. You also specify the source and target connections and the entities to include. Then, you configure the different phases in the retirement process and schedule the project.

1. In Data Archive, click **Workbench** > **Manage Retirement Projects**.

   The **Manage Retirement Projects** page appears.
2. Click **New Retirement Project**.

   The **Create or Edit a Retirement Project** page appears.
3. Enter the following information about the application:

   Application Name

   Required. Name of the application.

   Application Owner

   Optional. Name of the person responsible for the legacy application.

   Retirement Date

   Required. Date you plan to retire the application.

   Primary Contact

   Optional. Contact information of the application owner.

   Application Vendor Name

   Optional. Name of the application vendor.

   Vendor Contact

   Optional. Contact information of the application vendor.

   Vendor Phone

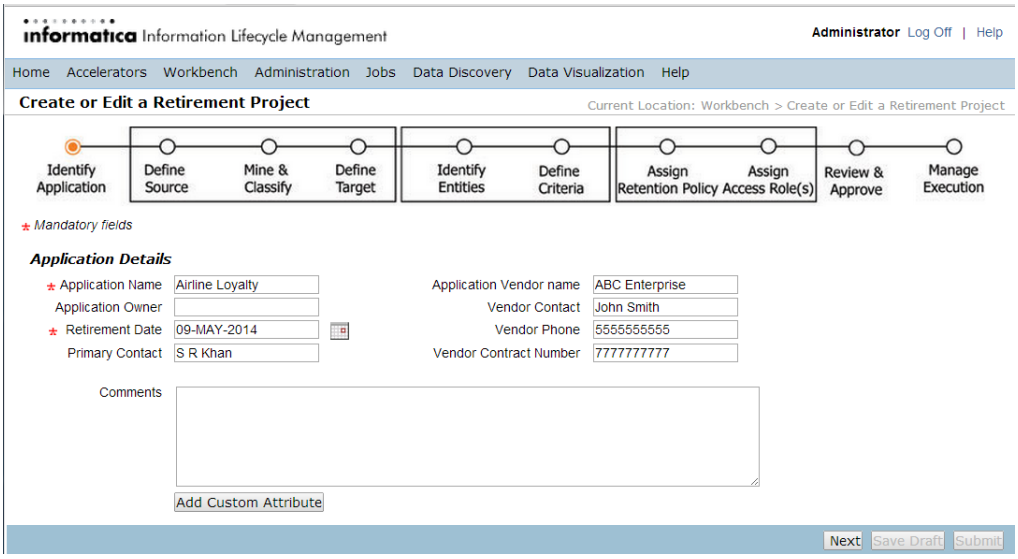
   Optional. Phone number of the application vendor.

   Vendor Contract Number

   Optional. The signed contract number of the application vendor.

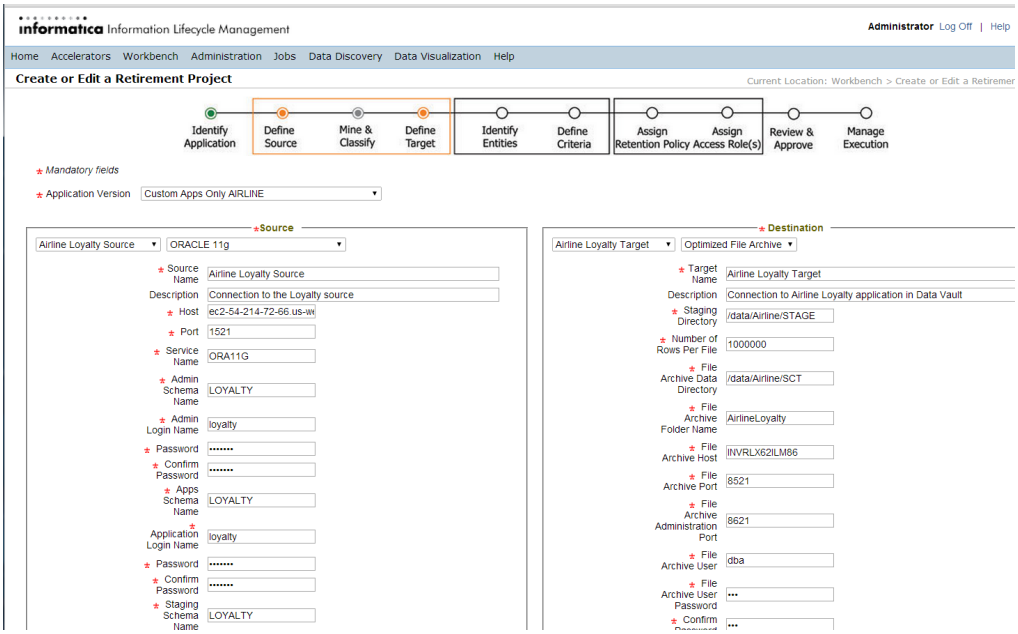
Comments

Optional. Additional information related to the retirement.



4. Optionally, to include additional application details, click the **Add Custom Attribute**. Enter a name and value for the parameter.
   For example, you can add an attribute for geographical location of source server.
5. Click **Next**.
6. Select the **Application Version**.

   You created the application version for the legacy application in Enterprise Data Manager.
7. Select the **Source** connection.
8. Select the **Target** connection.



9. Click **Next**.

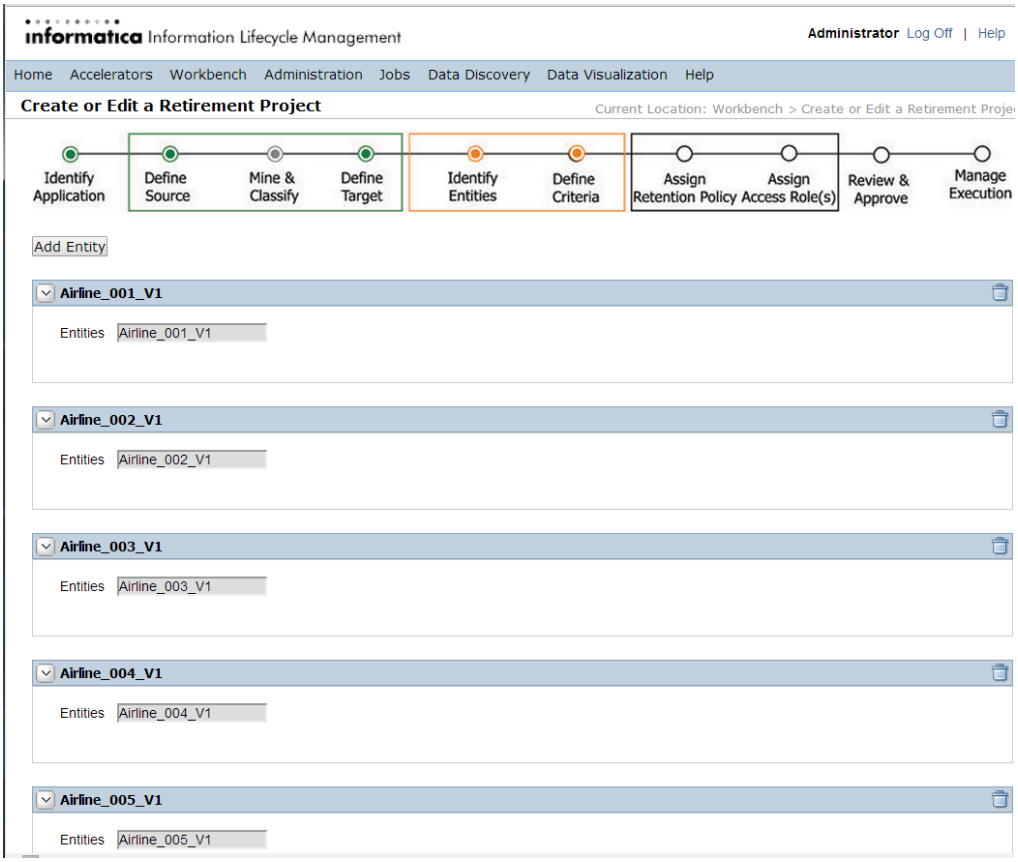
10. Click **Add Entity**.

    A window with a list of the retirement entities that you created in Enterprise Data Manager appears.

11. Press the SHIFT key and select all the entities.

12. Click **Select**.

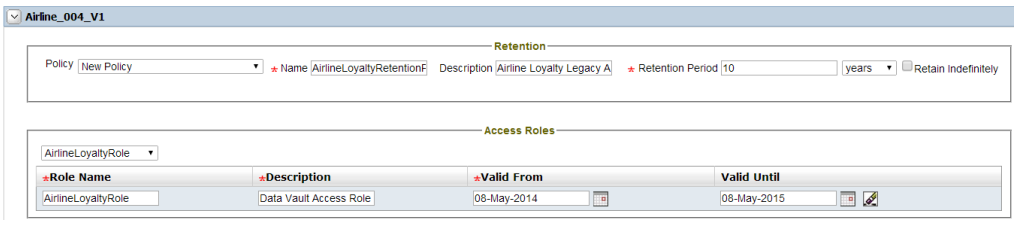
    The list of entities window closes. The list of entities appear on the **Create or Edit a Retirement Project** page.



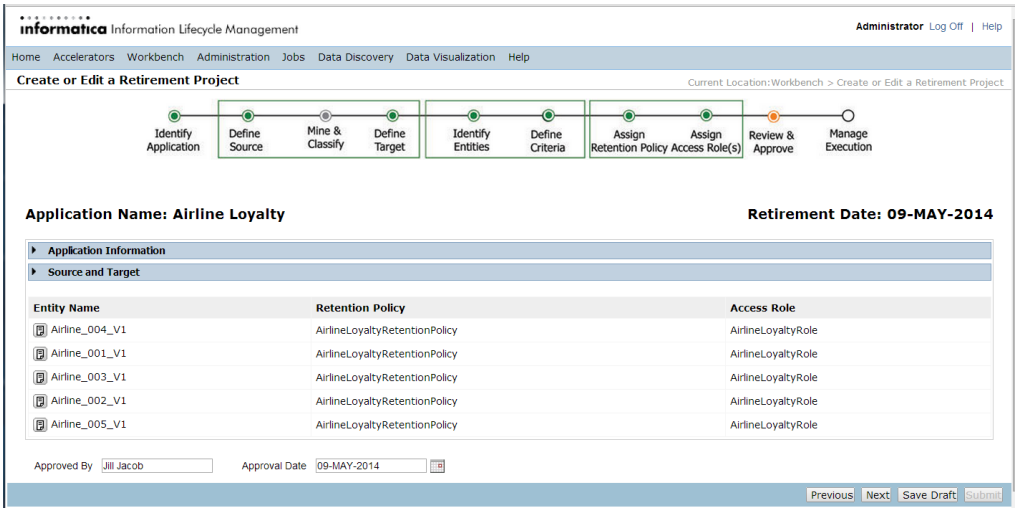
13. Click **Next**.

14. Optionally, in the **Retention** section, create or select a retention policy for each entity.

15. In the **Access Roles** section, assign the Data Vault access role that you created in "Step 7. Create a Data Vault Access Role" on page 20, to each entity.



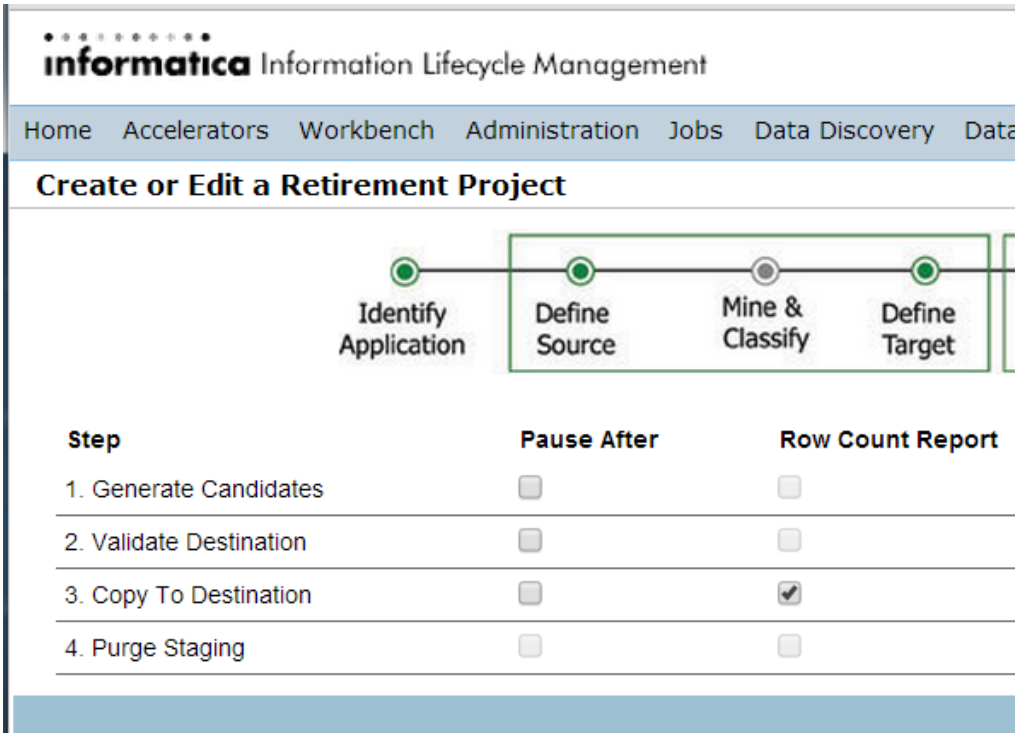
16. Click **Next**.
17. Optionally, enter the name of the person who approved the retirement and the date.



18. Click **Next**.

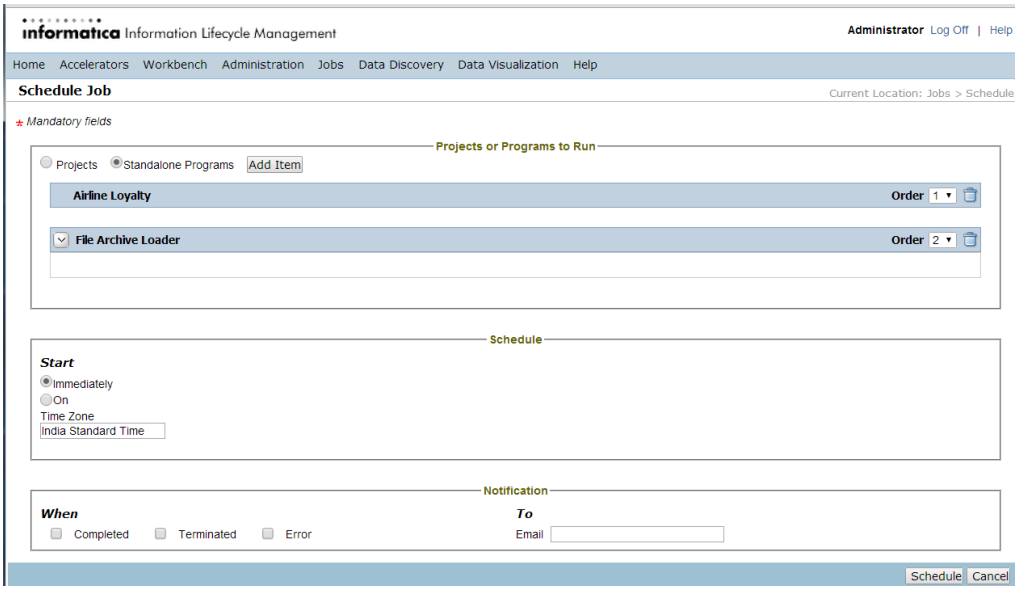
    The retirement steps appear.

19. Select the **Row Count Report** check box for the step **Copy to Destination**.



20. Select **Publish and Schedule**.

The **Schedule Job** page appears.



21. Select Standalone Programs and click **Add Item**.

    A list of jobs appears.

22. Select the **Create Archive Folder** job.

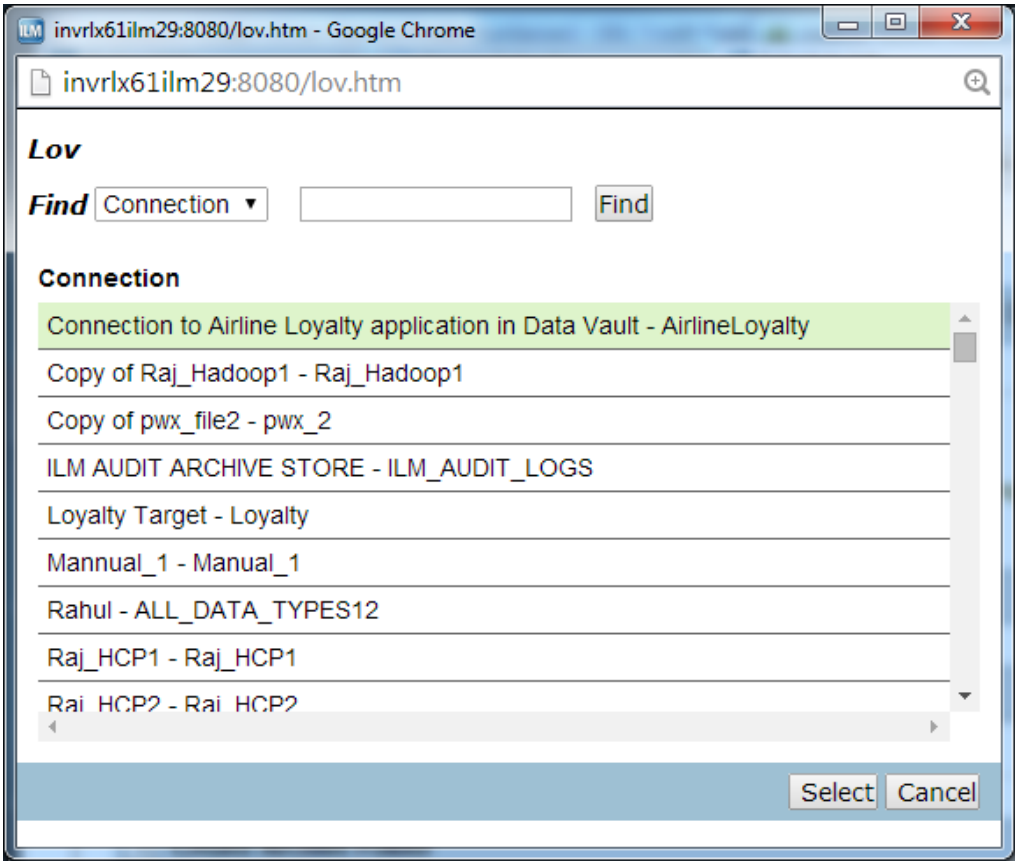
23. Click **Select**.

    The list of jobs closes. The **Create Archive Folder** job appears on the **Schedule Job** page.

24. Click the list of values icon next to the **Destination Repository**.

    A list of destination repositories appears.

25. Select the target repository you created for the application that you want to retire.



26. Click **Select**.

    The list of destination repositories closes.

27. Specify the following order for the jobs:

    a. Set order for the **Create Archive Folder** job to **1**.

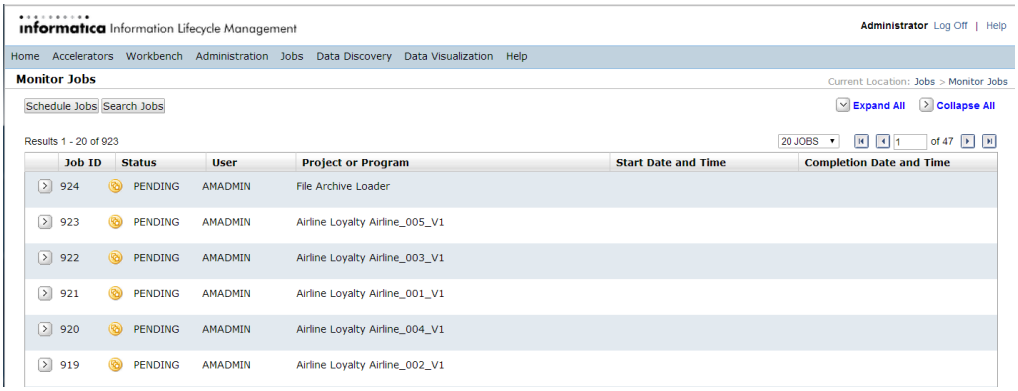
    b. Set order for the retirement project to **2**.

    c. Set order for the **Create Archive Folder** job to **3**.

28. To start the retirement process immediately, select **Immediately** in the **Schedule** section.

29. Click **Schedule**.

    The **Monitor Jobs** page appears with the status of the retirement jobs.

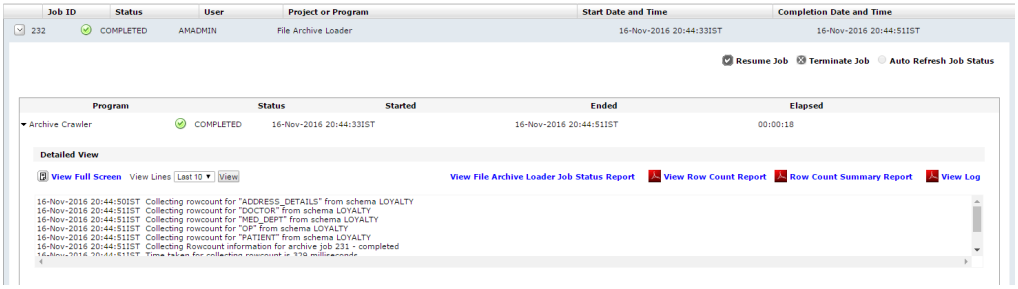


    When the job is complete, the job status changes to COMPLETED.

Step 10. Review the Row Count Report

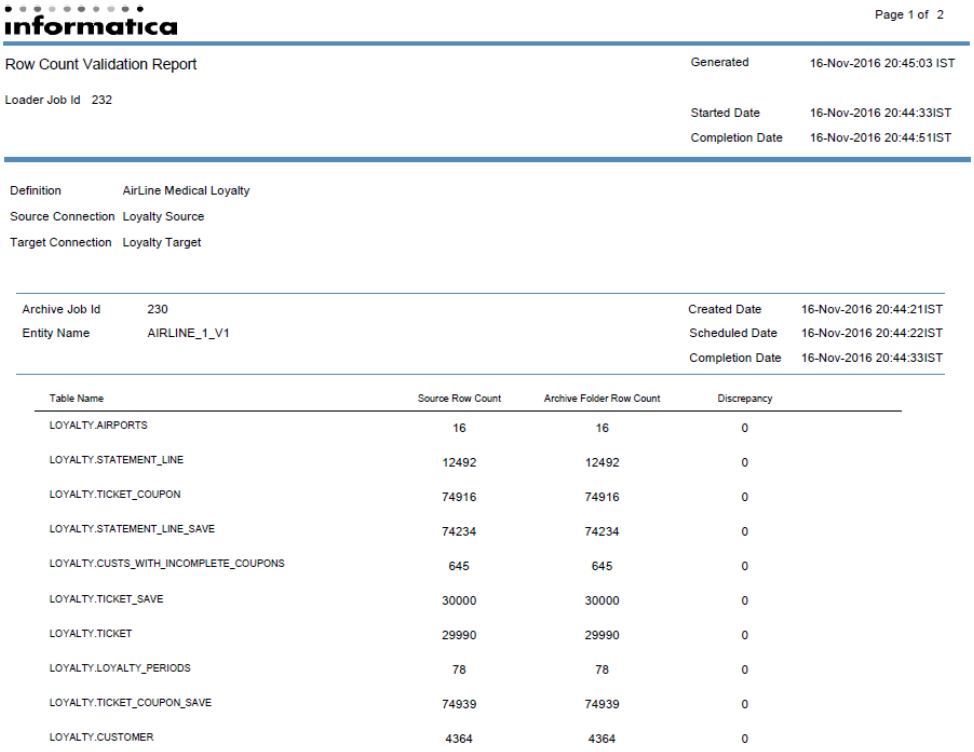
Review the row count report to verify that the Data Vault Loader job moved the data to Data Vault.

1. In Data Archive, click **Jobs** > **Monitor Jobs**.
2. Wait for the Data Vault Loader job to complete.
3. Expand the view for the Data Vault Loader job.
4. Expand the view for the Archive Crawler program.



5. Click **View Row Count Report**.

The row count report opens in PDF format.



informatica

Page 1 of 2

Row Count Validation Report

Generated 16-Nov-2016 20:45:03 IST

Loader Job Id 232

Started Date 16-Nov-2016 20:44:33IST

Completion Date 16-Nov-2016 20:44:51IST

Definition AirLine Medical Loyalty

Source Connection Loyalty Source

Target Connection Loyalty Target

Archive Job Id 230

Entity Name AIRLINE_1_V1

Created Date 16-Nov-2016 20:44:21IST

Scheduled Date 16-Nov-2016 20:44:22IST

Completion Date 16-Nov-2016 20:44:33IST

Table Name	Source Row Count	Archive Folder Row Count	Discrepancy
LOYALTY.AIRPORTS	16	16	0
LOYALTY.STATEMENT_LINE	12492	12492	0
LOYALTY.TICKET_COUPON	74916	74916	0
LOYALTY.STATEMENT_LINE_SAVE	74234	74234	0
LOYALTY.CUSTS_WITH_INCOMPLETE_COUPONS	645	645	0
LOYALTY.TICKET_SAVE	30000	30000	0
LOYALTY.TICKET	29990	29990	0
LOYALTY.LOYALTY_PERIODS	78	78	0
LOYALTY.TICKET_COUPON_SAVE	74939	74939	0
LOYALTY.CUSTOMER	4364	4364	0

Alternatively, you can click the **Row Count Summary** link to review the row count summary report.

6. For each table name, check the **Discrepancy** column.

   If the **Discrepancy** column is 0, then the Data Vault Loader job moved the data from the source to the Data Vault.

Step 11. Validate Data through the Data Discovery Portal

Use Browse Data from the Data Discovery portal to search and examine application data from specific tables in Data Vault.

When you use Browse Data, you specify the archive folder, entity, schema, and table you want to browse. You specify the table columns to display in the search results. You can use a `WHERE` clause to limit the search results. You can also specify the sort order of the search results. Before you run the search, you can preview the SQL statement that the Data Discovery portal uses to access the data.

1. Click **Data Discovery** > **Browse Data**.
2. Specify the following search parameters:

   Archive Folder

   Archive folder that contains the entity of the archived data that you want to search.

No. of Rows

Maximum amount of rows in the search results. Default is 100.

Entity

Entity that contains the table that you want to view archived data from.
Optional. Select the entity to limit the amount of tables shown in the list of values.

Schema

Schema that contains the table that you want to view.

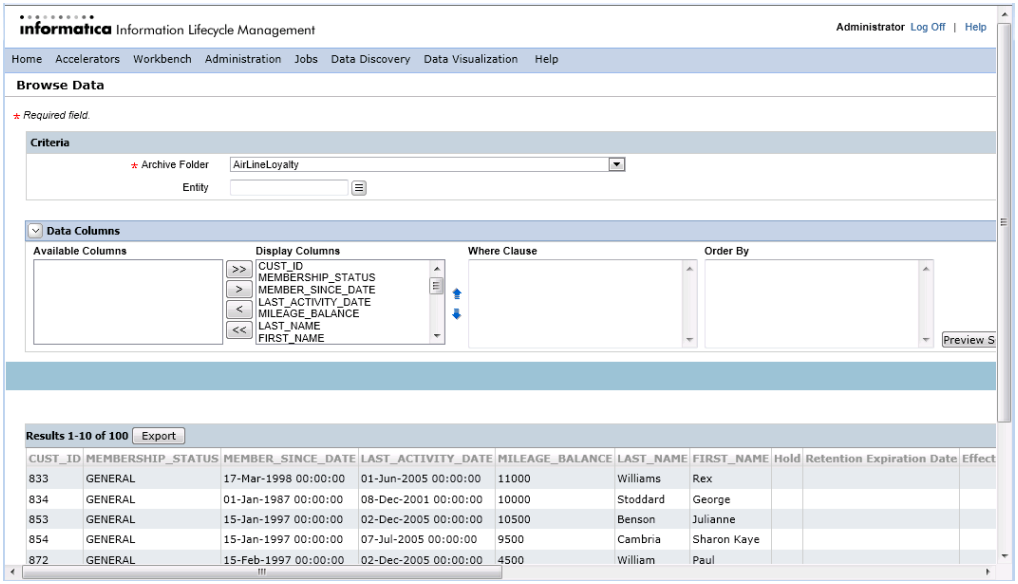
Table

Table that includes the archived data that you want to view.

After you select the table, Data Archive populates the **Available Columns** list.

3. Use the arrows to move the columns that you want to display to the **Display Columns** list.
4. Optionally, enter a `WHERE` clause in the **Where Clause** box to filter the search results.
5. Optionally, specify the sort order of the search results in the **Order By** box.
6. Optionally, click **Preview SQL** to view the generated SQL query.
7. Click **Search**.

   The records that match the search criteria appear.



8. Verify that the values for the records in Data Vault match the values for the same records on the source database.

After You Retire the Application

After you retire the application, you can access the data and enable compliance.

You can access the retired application in the following ways:

- Browse, search, and extract data through the Data Archive data discovery option.

- Query data with the Data Vault SQL Tool.
- Submit an SQL query through Data Visualization to create a report.
- Use a third-party tool to access data in Data Vault.

You can enable compliance in the following ways:

- Apply a retention policy to retain data for a required period of time.
- Apply a legal hold to protect data from modification.

Access to data in Data Vault is based on a system of users and the roles assigned to the user. A user must have the same role as the entity to access and enable compliance to the data in Data Vault.

You cannot edit or re-run a retirement project that has successfully completed.

INDEX

I