



## **erwin Data Modeler**

# **Automated Metadata Harvesting**

**Release 2019 R1**

## Legal Notices

This Documentation, which includes embedded help systems and electronically distributed materials (hereinafter referred to as the “Documentation”), is for your informational purposes only and is subject to change or withdrawal by erwin Inc. at any time. This Documentation is proprietary information of erwin Inc. and may not be copied, transferred, reproduced, disclosed, modified or duplicated, in whole or in part, without the prior written consent of erwin Inc.

If you are a licensed user of the software product(s) addressed in the Documentation, you may print or otherwise make available a reasonable number of copies of the Documentation for internal use by you and your employees in connection with that software, provided that all erwin Inc. copyright notices and legends are affixed to each reproduced copy.

The right to print or otherwise make available copies of the Documentation is limited to the period during which the applicable license for such software remains in full force and effect. Should the license terminate for any reason, it is your responsibility to certify in writing to erwin Inc. that all copies and partial copies of the Documentation have been returned to erwin Inc. or destroyed.

TO THE EXTENT PERMITTED BY APPLICABLE LAW, ERWIN INC. PROVIDES THIS DOCUMENTATION “AS IS” WITHOUT WARRANTY OF ANY KIND, INCLUDING WITHOUT LIMITATION, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT. IN NO EVENT WILL ERWIN INC. BE LIABLE TO YOU OR ANY THIRD PARTY FOR ANY LOSS OR DAMAGE, DIRECT OR INDIRECT, FROM THE USE OF THIS DOCUMENTATION, INCLUDING WITHOUT LIMITATION, LOST PROFITS, LOST INVESTMENT, BUSINESS INTERRUPTION, GOODWILL, OR LOST DATA, EVEN IF ERWIN INC. IS EXPRESSLY ADVISED IN ADVANCE OF THE POSSIBILITY OF SUCH LOSS OR DAMAGE.

The use of any software product referenced in the Documentation is governed by the applicable license agreement and such license agreement is not modified in any way by the terms of this notice.

The manufacturer of this Documentation is erwin Inc.

Provided with “Restricted Rights.” Use, duplication or disclosure by the United States Government is subject to the restrictions set forth in FAR Sections 12.212, 52.227-14, and 52.227-19 (c)(1) - (2) and DFARS Section 252.227-7014(b)(3), as applicable, or their successors.

Copyright © 2019 erwin Inc. All rights reserved. All trademarks, trade names, service marks, and logos referenced herein belong to their respective companies.

## Contact erwin

### Understanding your Support

Review [support maintenance programs and offerings](#).

### Registering for Support

Access the [erwin support](#) site and click Sign in to register for product support.

### Accessing Technical Support

For your convenience, erwin provides easy access to "One Stop" support for all editions of [erwin Data Modeler](#), and includes the following:

- Online and telephone contact information for technical assistance and customer services
- Information about user communities and forums
- Product and documentation downloads
- erwin Support policies and guidelines
- Other helpful resources appropriate for your product

For information about other erwin products, visit <http://erwin.com/products>.

### Provide Feedback

If you have comments or questions, or feedback about erwin product documentation, you can send a message to [techpubs@erwin.com](mailto:techpubs@erwin.com).

### erwin Data Modeler News and Events

Visit [www.erwin.com](http://www.erwin.com) to get up-to-date news, announcements, and events. View video demos and read up on customer success stories and articles by industry experts.

# Contents

---

<b>Legal Notices</b> .....	<b>2</b>
<b>Contents</b> .....	<b>4</b>
<b>Introduction</b> .....	<b>5</b>
Installing and Configuring Automated Metadata Harvesting (AMH) .....	6
AMHSettings.cfg File .....	12
Upgrade Considerations .....	13
Creating and Running Jobs .....	13
Creating or Importing an Application .....	13
Creating a Datastore .....	15
Setting a Reverse Engineering (RE) Option Set .....	17
Creating a Job .....	19
Running a Job .....	22
Viewing Job History .....	23
Deleting a Job .....	26
Best Practices .....	28
Viewing Mappings .....	29
Viewing Reports .....	31

## Introduction

Automated Metadata Harvesting (AMH) integrates erwin Data Modeler (DM) with erwin Data Governance (DG) and provides a way for Automatic Reverse Engineering (RE) for the databases used by your applications. It provides you with a way to govern all your applications. In other words, it keeps you informed about what data an application stores and how it is stored.

In short, the following sequence of events occurs:

1. A job is scheduled on erwin DG, telling the AMH utility which DB to reverse engineer.
2. The AMH utility processes the request serially as follows:
  - a. Checks the job queue for scheduled jobs
  - b. Based on the queue, reverse engineers the application
  - c. Creates the model with its name in the format, Application's Line Of Business.Application ID.DataStore name.TableOwner.erwin
  - d. If an NSM file is available, applies naming standards on the REDB model
  - e. Saves the model:
    - ◆ DG, Mart, and Web Portal
    - ◆ DG only
    - ◆ DG and Mart

The AMH utility consists of two Windows services:

### **erwinDG Queue Service**

This service communicates with the erwin Data Governance (DG) platform to look for availability of new jobs. It creates a queue of jobs based on the information received from the erwin DG platform.

### **erwinDG DM Service**

This service communicates with the ErwinDG Queue Service to check for any jobs in the queue and executes them.

This section contains the following topics

## Installing and Configuring Automated Metadata Harvesting (AMH)

The AMH utility is one of the components in the erwin Data Modeler (DM) installer. Installing AMH installs erwinDG Queue Service and erwinDG DM Service on your environment. Also, the installation creates a Config folder at C:\Program Files\erwin\Data Modeler r9\, which contains the following:

1. **AMHSettings.cfg**: Stores AMH services-related configuration. For more information, read the [AMHSettings.cfg File](#) topic.
2. **Mart.cfg**: Stores erwin Mart-related configuration
3. **Platform.cfg**: Stores erwin Data Governance (DG) platform-related configuration
4. **MartWP.cfg**: Stores erwin Mart Web Portal-related configuration
5. **Databases.csv**: Stores database connectivity-related configuration, used for reverse engineering  
**Note:** Table Owner, Mart Folder, RE Option Set, and Model Target and Use DBA Views are handled by the DG Server instead of Databases.csv. Also, Table Owner and Mart Folder's size limit is 255.
6. **REOptionSet folder**: Contains three reverse engineering option sets for each database; custom, all, and default

**Note:** At a later point in time, if you wish to change the settings or update any of the above files manually, for the changes to take effect, run the following commands in Windows command line utility (in erwin home directory, with administrator privileges) and then restart AMH Windows services (erwinDG Queue Service and erwinDG DM Service):

- **AMHSettings.cfg:**

erwinDMAPI.exe -Settings

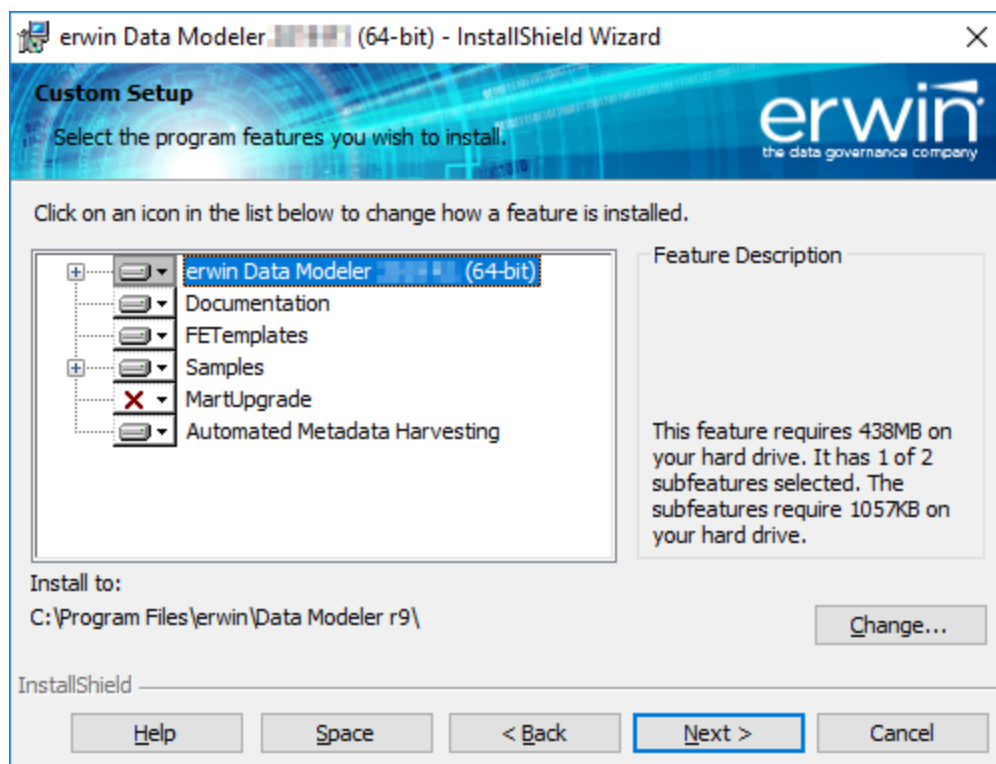
- **Mart.cfg / Platform.cfg / Databases.csv:**

erwinDMAPI.exe -encrypt

Other than the above files and folders, another file, DM\_AMH\_Trace.log, is created at C:\Program Files\erwin\Data Modeler r9\Logs. This file logs information about the actions performed by AMH.

Follow these steps to install AMH on your environment:

1. On the erwin DM installer, select Automated Metadata Harvesting and click Next.



2. Enter values for erwin Mart Server settings and click Next.

erwin Data Modeler (64-bit) - InstallShield Wizard

Enter information in the fields below  
The installer will use this information in the subsequent steps.

erwin  
the data governance company

Mart

Server Name: localhost

Port: 18170 ☐ Use SSL (https)

IIS Port: 80 ☐ Use IIS

User Name:

Password:

InstallShield

< Back Next > Cancel

Field descriptions:

- **Server Name:** Indicates the name of the web server where you have installed erwin Mart. The default value is localhost.  
**Note:** Do not use an IP address as server name to avoid the naked domain problem.
- **Port:** Indicates the port number to access the web server. The default value is 18170.
- **Use SSL (https):** Indicates whether you want to connect to the mart through a secured connection. By default, this check box is not selected.
- **IIS Port:** Indicates the IIS port number. The default value is set to 80.



- **Use IIS:** Indicates whether you want to use IIS to connect to the Mart. By default, this check box is not selected.
- **User Name:** Indicates your erwin Mart user name.
- **Password:** Indicates your password to connect to erwin Mart.

The settings you make here are saved to a configuration file, Mart.cfg, at C:\Program Files\erwin\Data Modeler r9\Config.

3. Enter values for erwin DG platform settings and click Next.

Field descriptions:

- **Server Name:** Indicates the name of the web server of the erwin DG platform. This format is <client Name>.<web server>. For example, erwin-poc-amh.my-erwin.io. By default, the web server is set to myerwin.io. Client Name indicates the name of client page for each customer. For example, erwin-poc-amh.

- **Port:** Indicates the port number to access the web server. The default value is 443.
- **Use SSL (https):** Indicates whether you want to connect to the erwin DG platform through a secured connection. By default, this check box is selected.
- **User Name:** Indicates your erwin DG platform user name.
- **Password:** Indicates your password to connect to erwin DG platform.

The settings you make here are saved to a configuration file, Platform.cfg, at C:\Program Files\erwin\Data Modeler r9\Config.

4. Enter values for erwin Mart Web Portal settings and click Next.

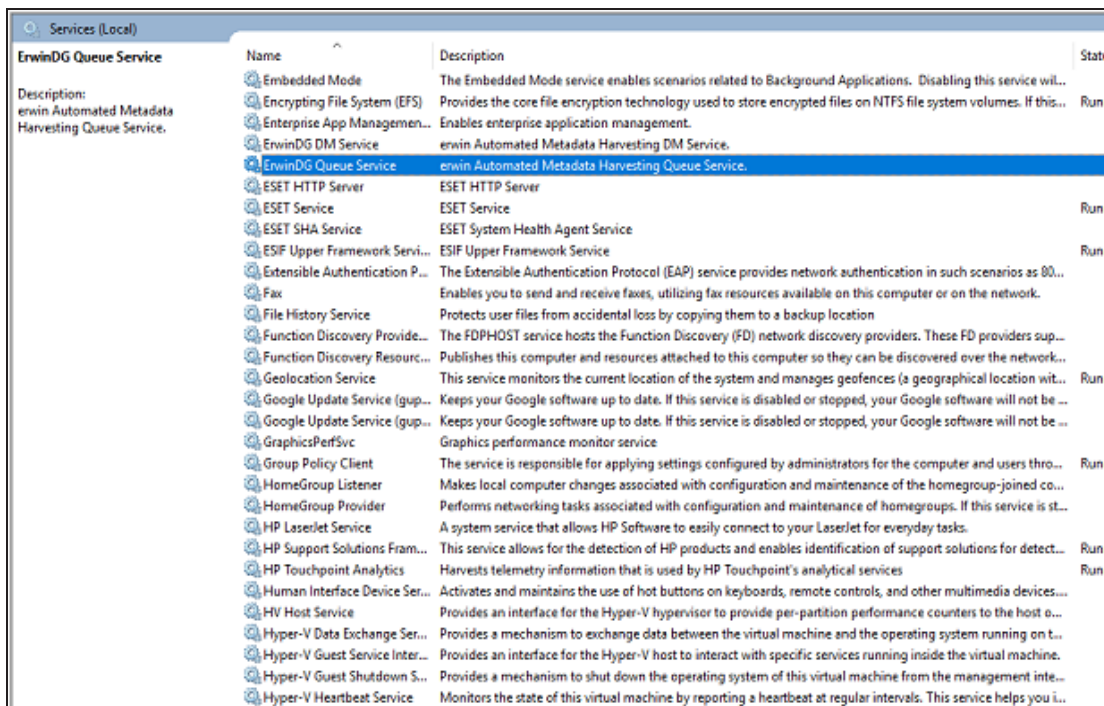
Field descriptions:

- **Server Name:** Indicates the machine name of the Web Portal server running on the machine. By default, it is set to localhost.

- **Version:** Indicates the Web Portal server's version. By default, it is set to 2018 R1.
- **Port:** Indicates the port number to access the web portal. The default value is 11680.
- **Use SSL (https):** Indicates whether you want to connect to the erwin Mart Web Portal through a secured connection. By default, this check box is not selected.
- **Published Configuration Name:** Indicates the name of the folder under which models should be published on the Web Portal.

The settings you make here are saved to a configuration file, MartWP.cfg, at C:\Program Files\erwin\Data Modeler r9\Config.

5. Click Next and then click Install. AMH (ErwinDG Queue Service and ErwinDG DM Service) will be installed on your environment.



Also, a Databases.csv file that contains all the information that you need for reverse engineering is created at C:\Program Files\erwin\Data Modeler r9\Config.

## AMHSettings.cfg File

The AMHSettings.cfg file contains the following configuration:

- **DG Queue Interval Time:** Indicates the interval in which the ErwinDG Queue Service asks erwin DG server for availability of new jobs. The default value is 5000 milliseconds (5 seconds).
- **DM Interval Time:** Indicates the interval in which the ErwinDG DM Service reads the job queue to check for new jobs. The default value is 10000 milliseconds (10 seconds).
- **Model Name With Time:** Indicates whether you want to save the reverse engineered models with a name that includes model name and the timestamp. The default value is No.
- **DM Instance:** Indicates the ID of your erwin DM object for which you want to run reverse engineering. You can find this on the erwin DG Platform. The default value is set to 0, which creates a new erwin DM object with your localhost name, on erwin DG Platform.
- **TraceLogLevel:**
  - **Info:** General activity messages, including error messages. This is the default value.
  - **Verbose:** Info and detailed messages
  - **Debug:** Verbose, developer-specific messages, erwin DG Queue Trace message (Event log viewer displays this message)
  - **Detail:** Debug and detailed messages (For example, repeated general header messages)

The AMHSettings.cfg file is saved at C:\Program Files\erwin\Data Modeler r9\Config.

## Upgrade Considerations

While upgrading from erwin Data Modeler (DM) 2018R1 onward to the latest version, ensure that you follow this:

1. Back up the following folders:
  - Config (C:\Program Files\erwin\Data Modeler r9\Config)
  - Logs (C:\Program Files\erwin\Data Modeler r9\Logs)
2. Uninstall erwin DM (2018R1) and then install erwin DM (2019R1 or latest).
3. (Optional) Restore the Config and Logs folders.

## Creating and Running Jobs

Creating jobs is where erwin Data Governance (DG) platform comes into picture. When you install erwin Data Modeler (DM) with Automated Metadata Harvesting (AMH), an object of your erwin DM installation is created on the erwin DG platform with your machine name as the erwin DM object name.

Use AMH for reverse engineering requires you to go through each of the following steps:


1. [Creating or importing an application](#)
2. [Creating a datastore](#)
3. [Setting a Reverse Engineering \(RE\) option set](#)
4. [Creating a job](#)
5. [Running a job](#)

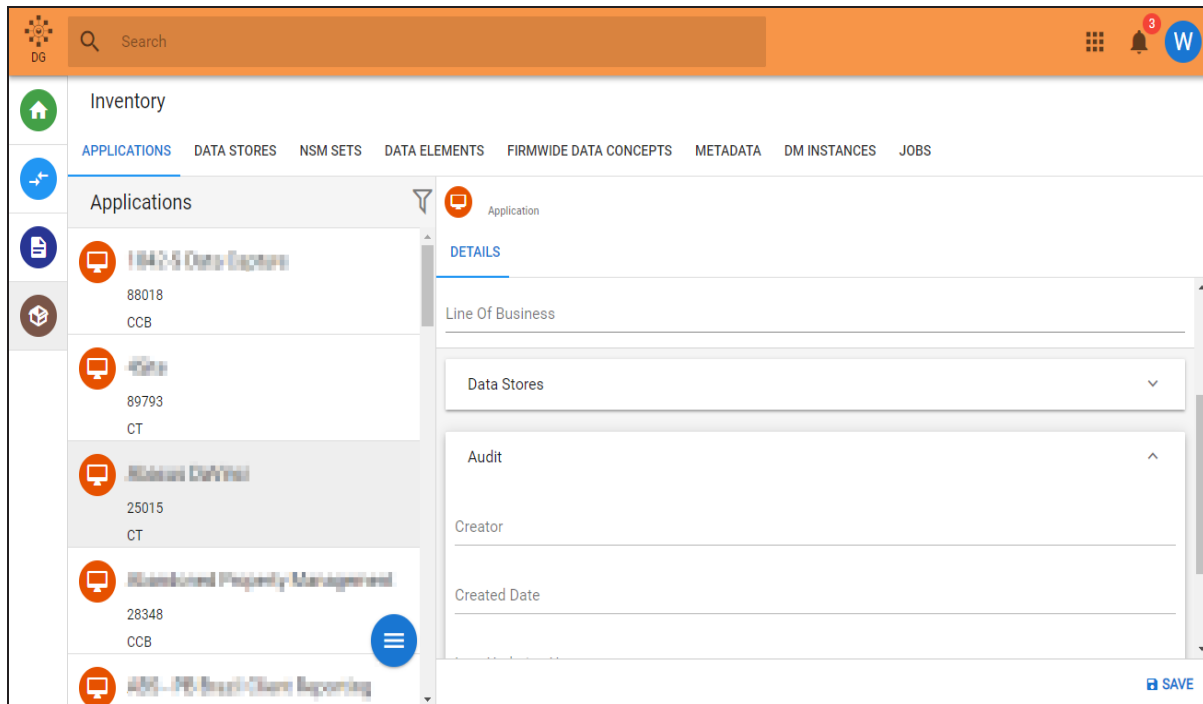
## Creating or Importing an Application

To be able to create a reverse engineering job, you first need to have your application available on the erwin DG platform. Then, you can create and associate a datastore to it.

To create or import an application, follow these steps:

1. Log on to erwin DG platform.

2. On the navigation bar on the left, click .  
This opens the Inventory page.



3. Click the Applications tab.

This opens a list of available applications and their corresponding information.

4. Click .

To create an application:

a. Click .


A new application instance is opened

b. Enter the Name, Description, Application ID, and Line of Business.

- c. Expand each tab and select / enter values of each parameter on the tabs. Set the existing datastores associated with the application. Also, enter audit information.
- d. Click Save.  
An application is created.

To import an application:




- a. Click .  
The import pane is opened.
- b. Select the type of file that you want to import.
- c. Depending on the type that you select, set/select other options.
- d. Drag and drop or click the Click to Upload button to select the file that you want to import.
- e. Click Import.

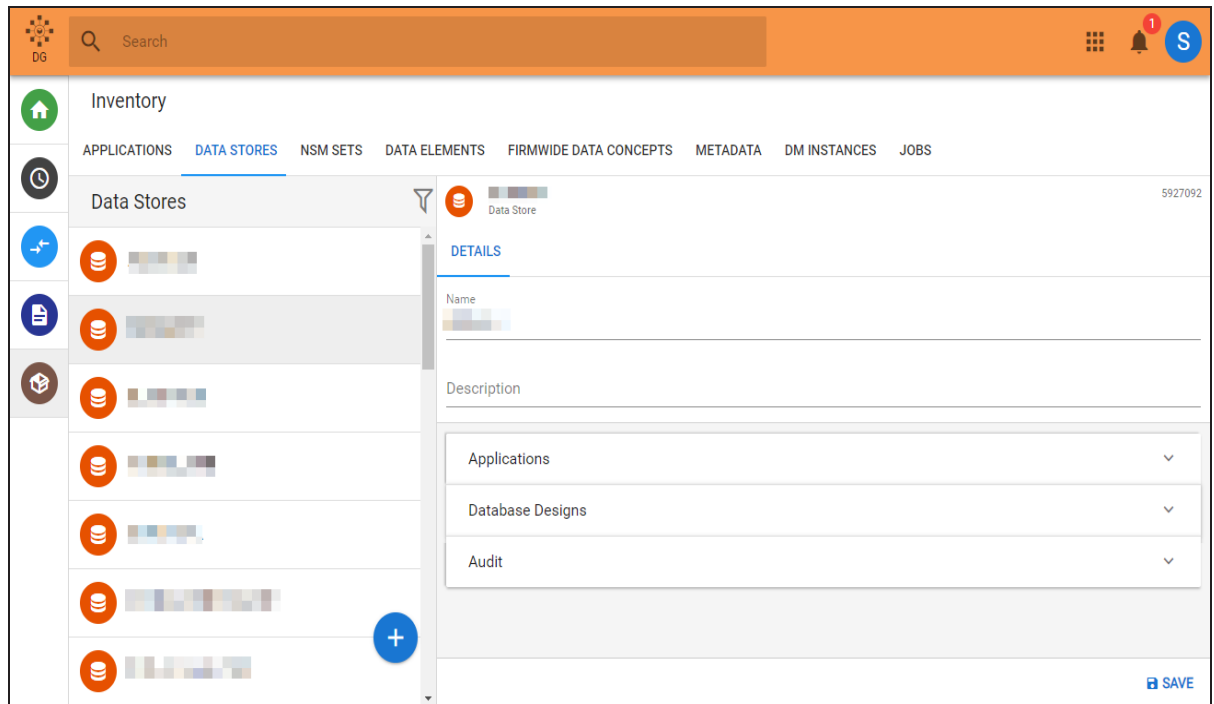
## Creating a Datastore

To be able to create a reverse engineering job, you first need to create a data store and associate it with your application. To create a datastore, follow these steps:

1. Log on to erwin DG platform.



2. On the navigation bar on the left, click .  
This opens the Inventory page.



3. Click the Data Stores tab.

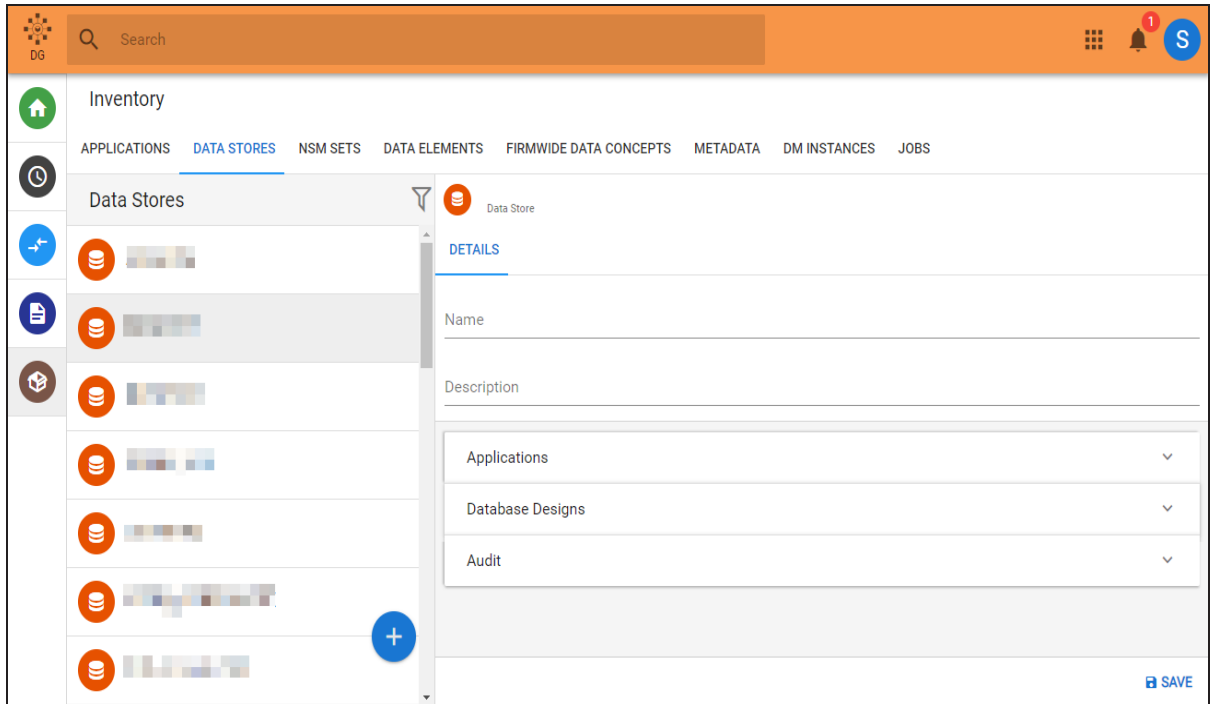
This opens a list of data stores available and their corresponding information.



4. Click .

A new data store instance is opened






5. Enter the Name and Description.
6. Expand each tab and select / enter values of each parameter on the tabs. Set the applications associated with the datastore. Also, enter audit information.  
**Note:** Database Designs are created by the reverse engineering process. You do not have to set it.
7. Click Save.  
A data store is created.

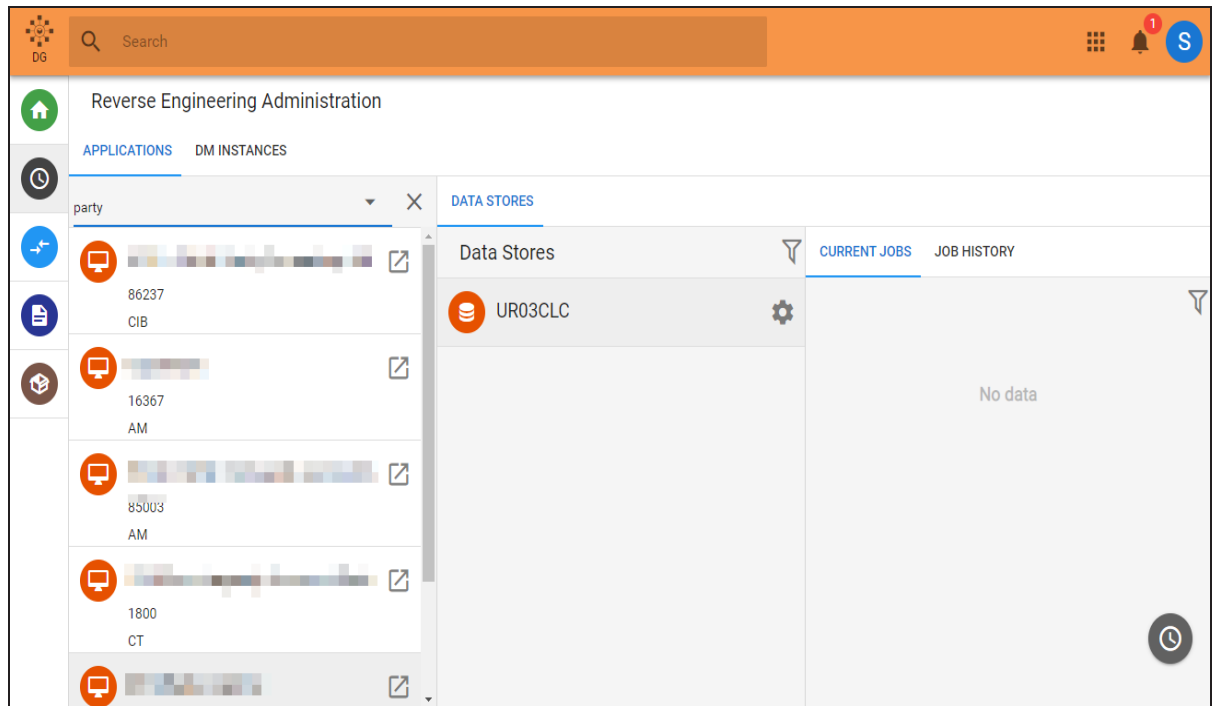
## Setting a Reverse Engineering (RE) Option Set

Once you have a data store in place, set an Reverse Engineering (RE) option set for it. To set an RE option set, follow these steps:

1. Log on to erwin DG platform.




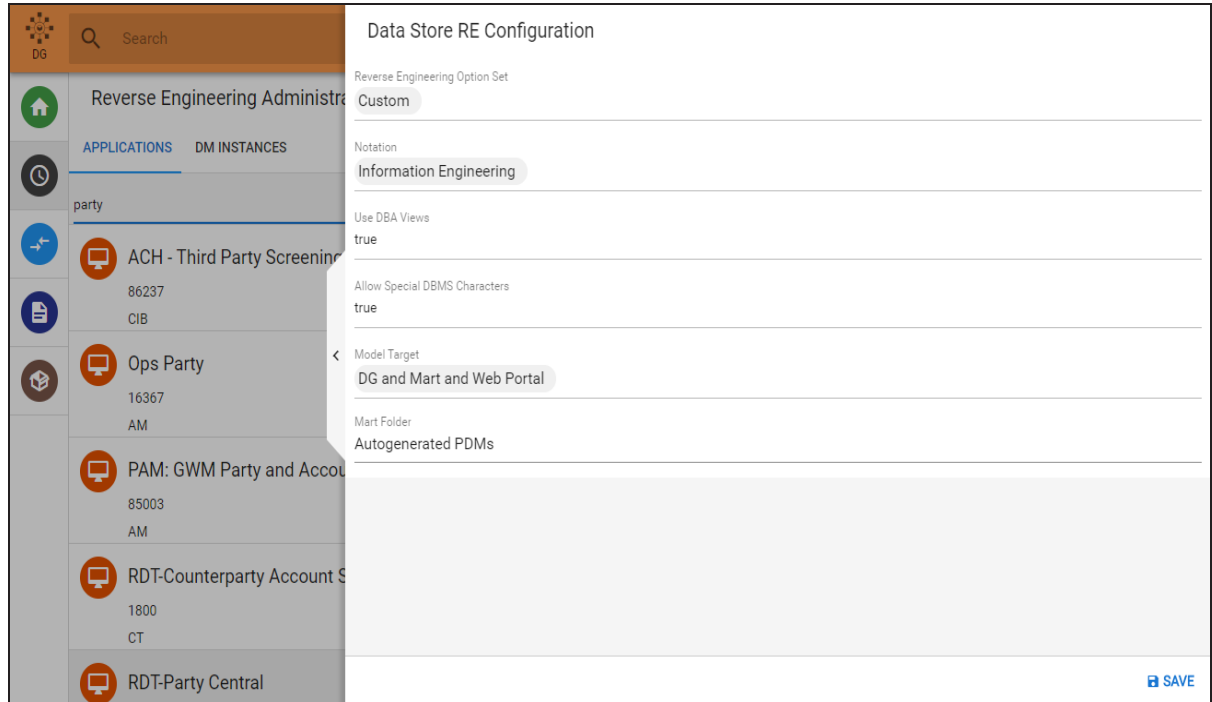
2. On the navigation bar on the left, click .  
This opens the Reverse Engineering Administration page.



3. Select your application.

The list of associated data stores is displayed in the Data Stores pane.


4. Click  to open your data store's RE configuration.



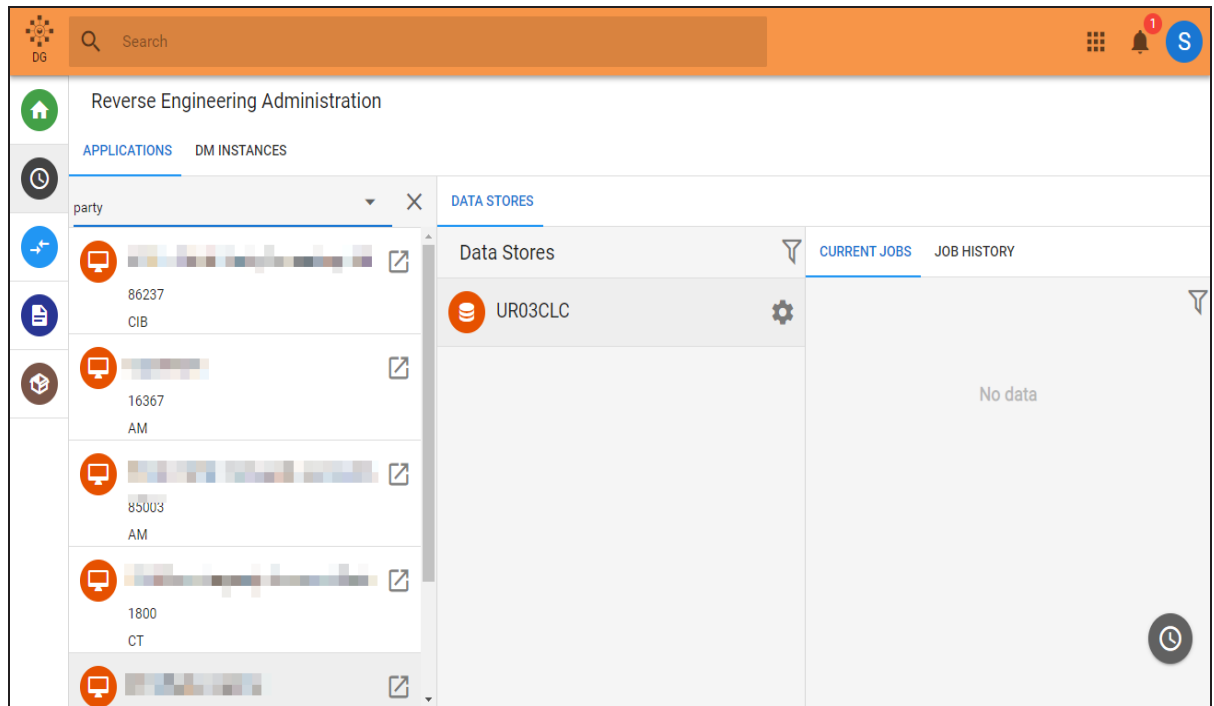
5. Select the Reverse Engineering Option Set, Notation, and other options.
6. Click Save.
- The RE option set for your data store is set.

## Creating a Job


After you have created your data store and its RE option set, to create a reverse engineering job, do the following:

1. Log on to erwin DG platform.
2. On the navigation bar on the left, click .

This opens the Reverse Engineering Administration page. A list of applications is displayed.



3. Select your application.  
A list of associated data stores is displayed.

4. Select the data store for which you want to create a job and click .  
The Create Job dialog is displayed.

Create Job

Name

Schedule Window Start

Schedule Window End

Table Owner

Application

erwin DM Instance

SAVE

5. Enter a Name.
6. Select Schedule Window Start and Schedule Window End to define the time window in which the job should run.  
Use the following values (indicative) as a guideline:
  - For a non-Oracle database, with around 5000 tables, a reverse engineering job would take around 30-45 minutes.
  - For an Oracle database, with around 5000 tables, a reverse engineering job would take around 3-4 hours.
7. Enter a Table Owner.
8. Select an application and an erwin DM Instance.
9. Click Save.

A new job is created and added to the list of jobs for your application.

## Running a Job

Once you have scheduled a job, ensure that the Automated Metadata Harvesting (AMH) services are running.

By default, the AMH services keep running in the background. Once the scheduled job runs, the reverse engineered models are saved locally, on erwin DM Mart, and on the erwin Data Governance Platform.

You can view the status of your job in the DM\_AMH\_Trace.log file, located at C:\Program Files\erwin\Data Modeler r9\Logs. You can also view the status on the erwin Data Governance platform at Data Usage > erwin DM Instances > <<your erwin DM object>> > Job.

In case you must run the services manually, run them in following order:

1. erwinDG Queue Service
2. erwinDG DM Service

To run these services manually, do the following:

1. Press the Windows and R (Windows+R) keys.  
This opens the Run dialog box.
2. Type services.msc and click OK.  
This opens the Windows Service Manager.
3. Locate the erwinDG Queue Service in the list of services, right-click and click Start.
4. Similarly, locate the erwinDG DM Service, right-click and click Start.  
This will run the reverse engineering job that you scheduled.

**Note:** Jobs are run serially. Hence, schedule jobs in a way that the schedule window is reasonable. Ensure that you consider the DB, its size, and the approximate job duration of the current jobs and then schedule a new job accordingly.

For example, consider the following scenario.

Job	Serial Number	Scheduled Time	Actual Required Time (Approximately)
A	1	2 hours	2 hours

B	2	4 hours	6 hours
C	3	30 minutes	30 minutes

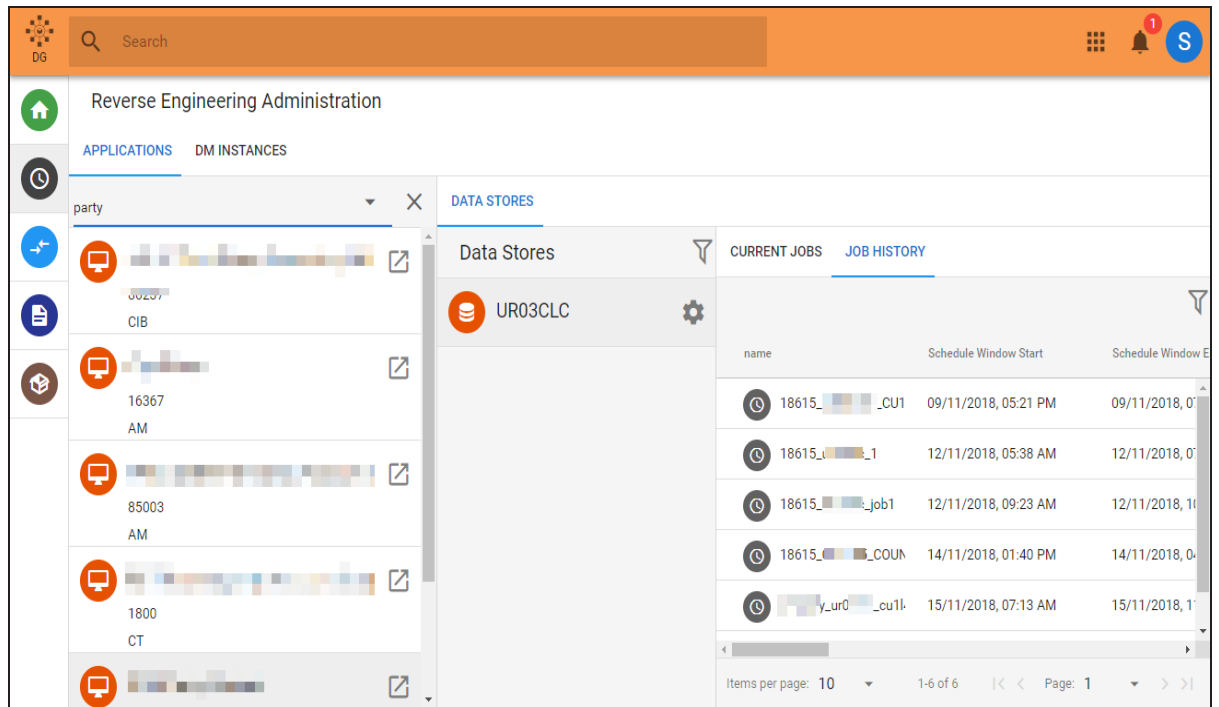
- Job A will run its course and complete in the scheduled window.
- Job B, according to you must complete in four hours. However, it takes 6 hours and keeps running beyond the scheduled window.
- Job C does not run and remains in scheduled state (becomes dormant) because job B kept running beyond the scheduled window and overlapped job C's window.  
To run job C, you need to reschedule it.

## Viewing Job History

For each data store, you can view its job history. To view the history, do the following:

1. On the Reverse Engineering Administration page,
2. Select your application.  
A list of associated data stores is displayed.
3. Select the data store for which you want to view the job history.  
A list of current jobs and job history is displayed.


4. Select the Job History tab.  
Job history for your data store is displayed.



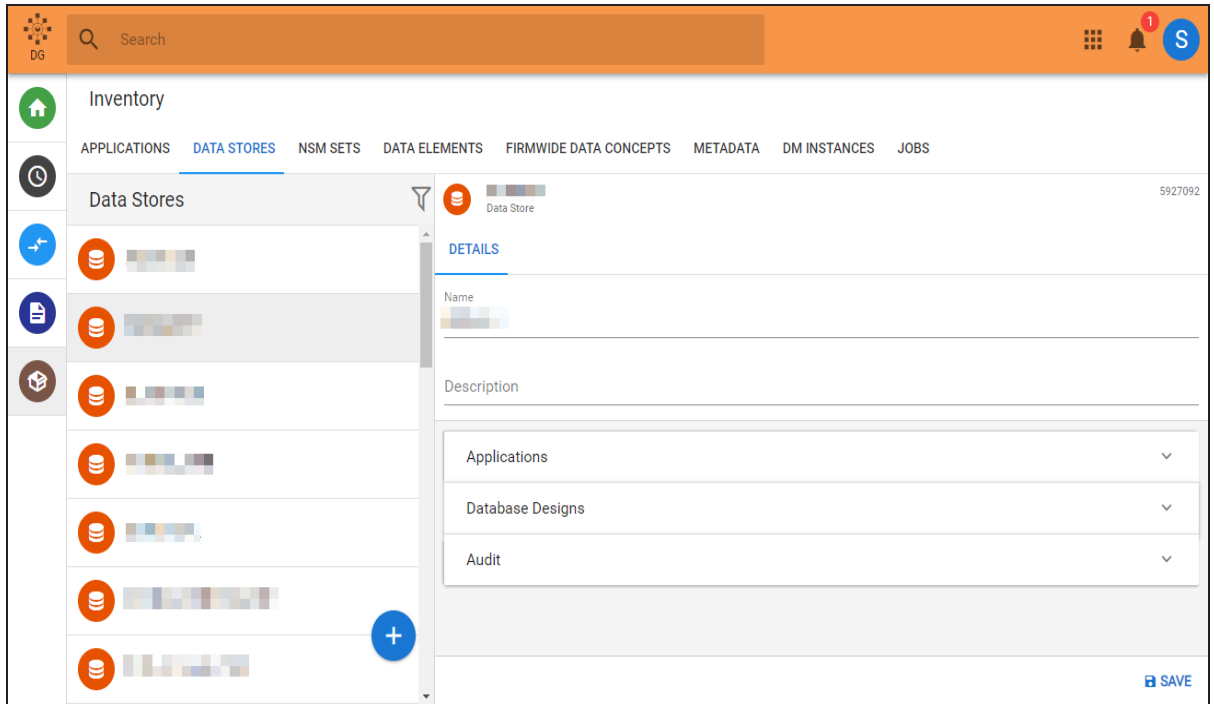
Alternately, you can view all the jobs on your account. To view all the jobs, do the following:

1. Log on to erwin DG platform.



2. On the navigation bar on the left, click .  
This opens the Inventory page.

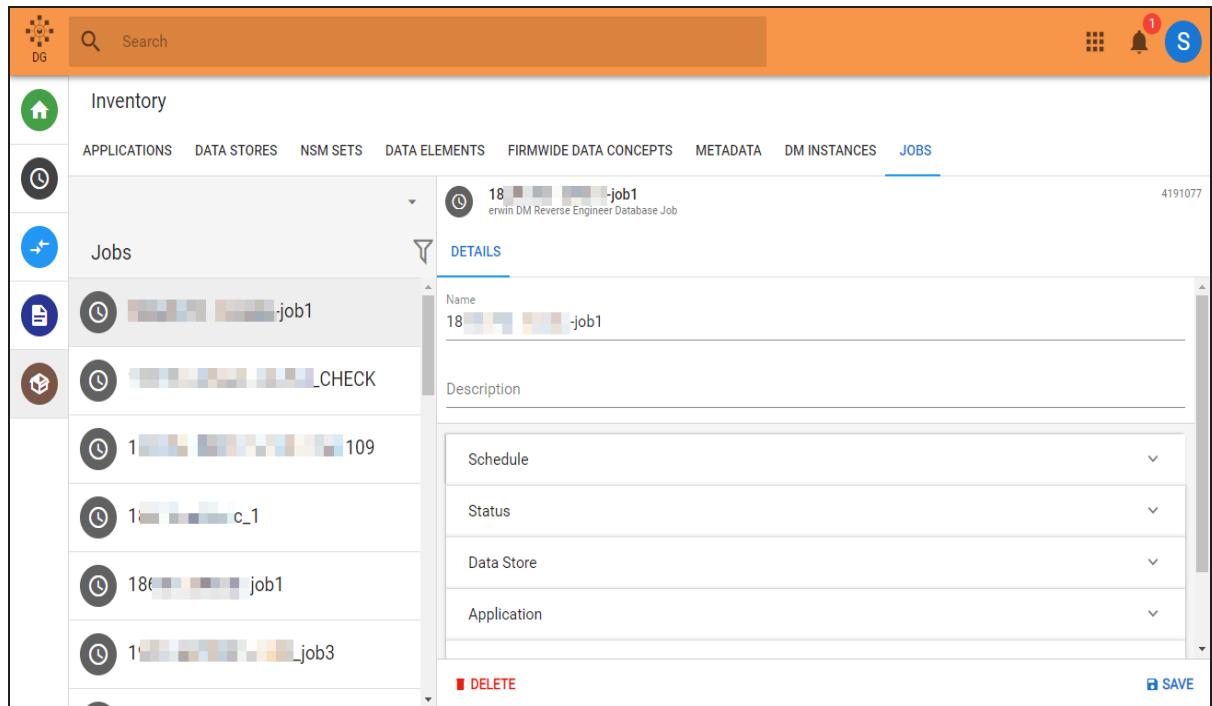




3. Click the Jobs tab.

A list of all jobs on your account is displayed. You can filter the list to view current or

completed jobs.



4. Select a job to view its details, such as its schedule, status, data store, application, DM object, and so on.


## Deleting a Job

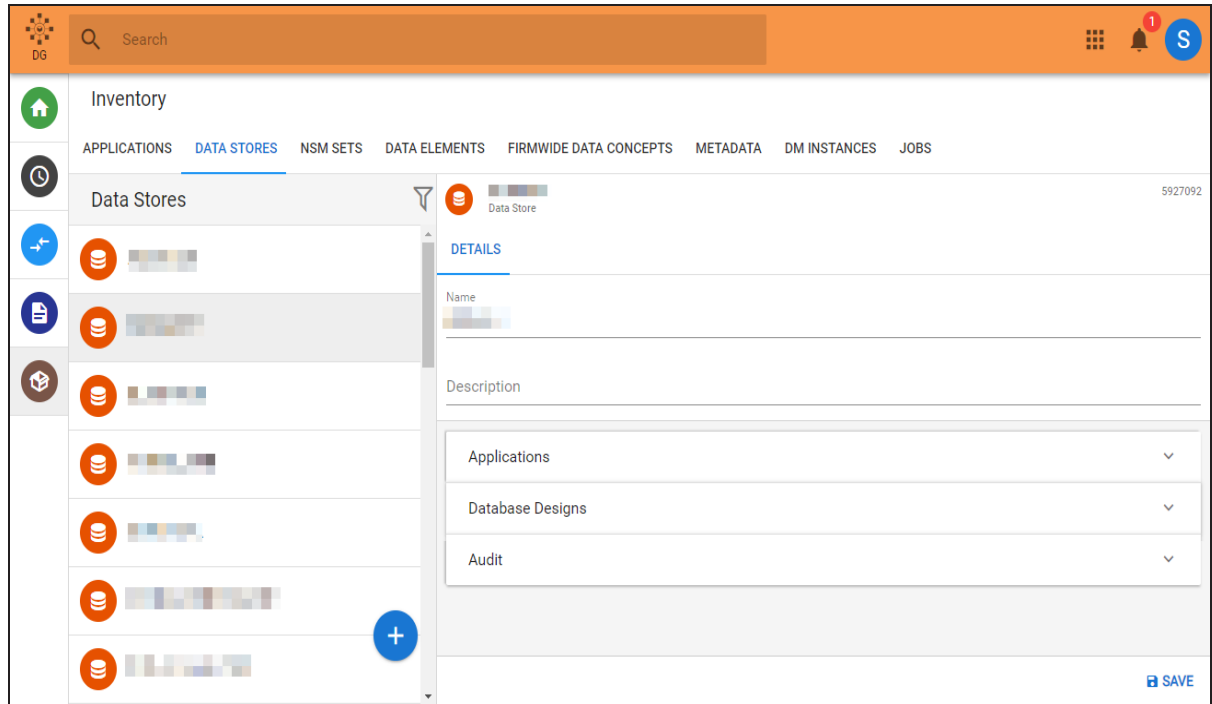
When you delete a job, the element representing your job is deleted. Consider the following cases while deleting a job:

- If the job is in queue, to be processed, it will be deleted from the queue and will not run.
- If the job is running:
  - If you have direct access to erwin Data Modeler (DM), you can stop erwin DM process to avoid the model from being saved anywhere.
  - If you do not have direct access to erwin DM, the reverse engineering (RE) process will continue and the model will be saved to the erwin Mart and erwin Web Portal. It will not be saved to erwin Data Governance (DG) platform.

To delete a job, do the following:

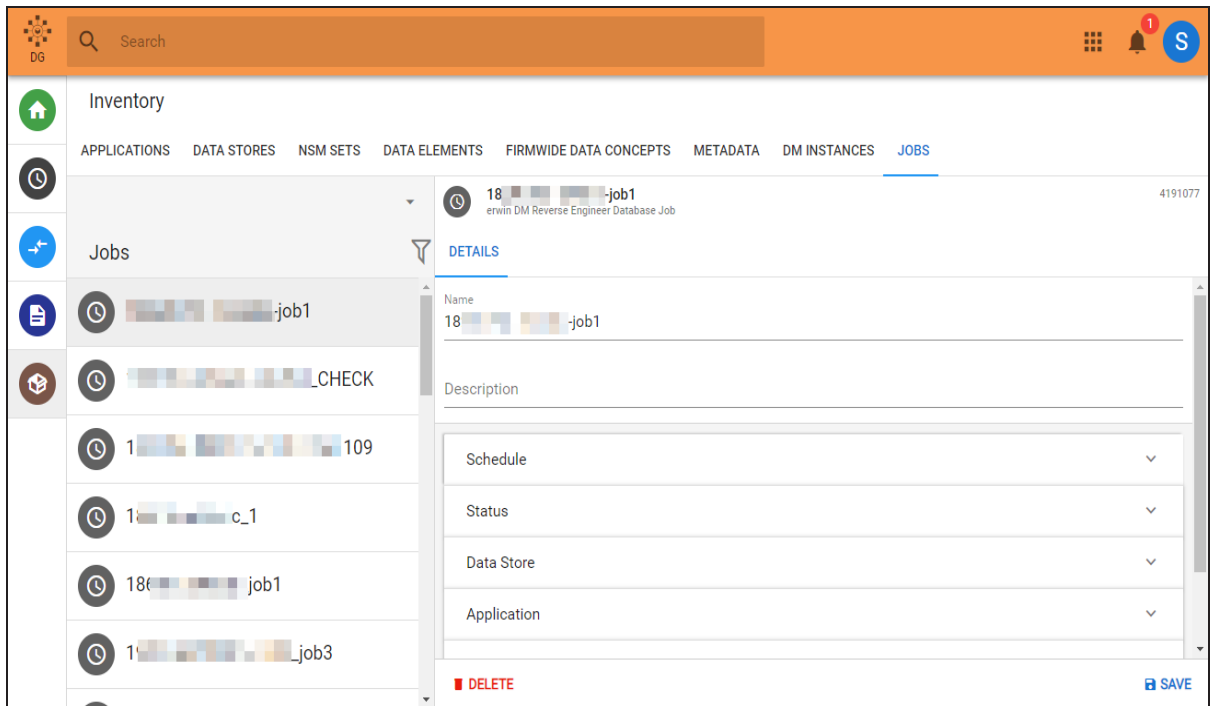
1. Log on to erwin DG platform.

2. On the navigation bar on the left, click .  
This opens the Inventory page.



3. Click the Jobs tab.  
A list of all jobs on your account is displayed. You can filter the list to view current or

completed jobs.



4. Select a job and click Delete.

## Best Practices


Consider the following while scheduling and running jobs:

- Ensure that you select all the parameters pertaining to job scheduling correctly. Else, a job might not run at all, even though you are able to schedule it.
- If a reverse engineering job executed in less than a minute, it means that either the table owners were not evident or the job could not find the right parameters.
- Ensure that you do not set start time  $\geq$  finish time. Although the job will be scheduled, it will not run.

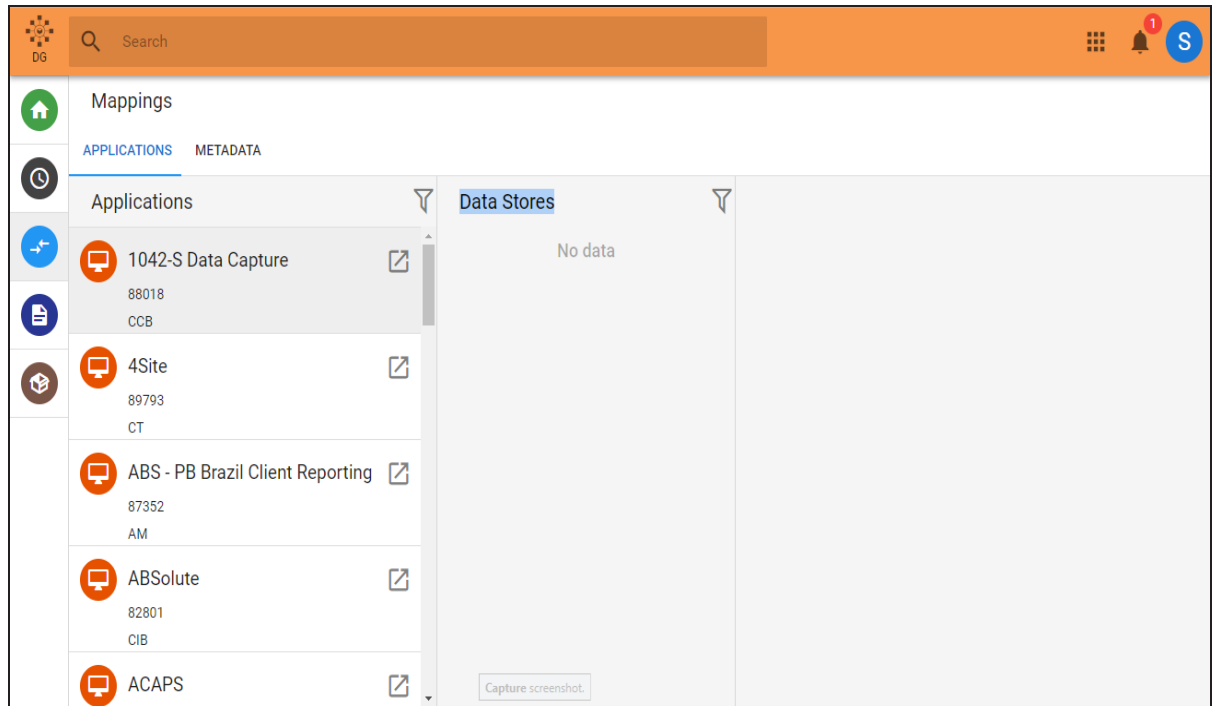
## Viewing Mappings

Mappings are generated after a reverse engineering job is completed successfully. A mapping provides information, such as data models associated with a data source, the physical and logical metadata, and the ER diagram. Using the Mappings page, you can map Columns to Data Elements and tables to Firmwide Data Concepts using the Expanded Logical Name (ELN).

To view a mapping, do the following:

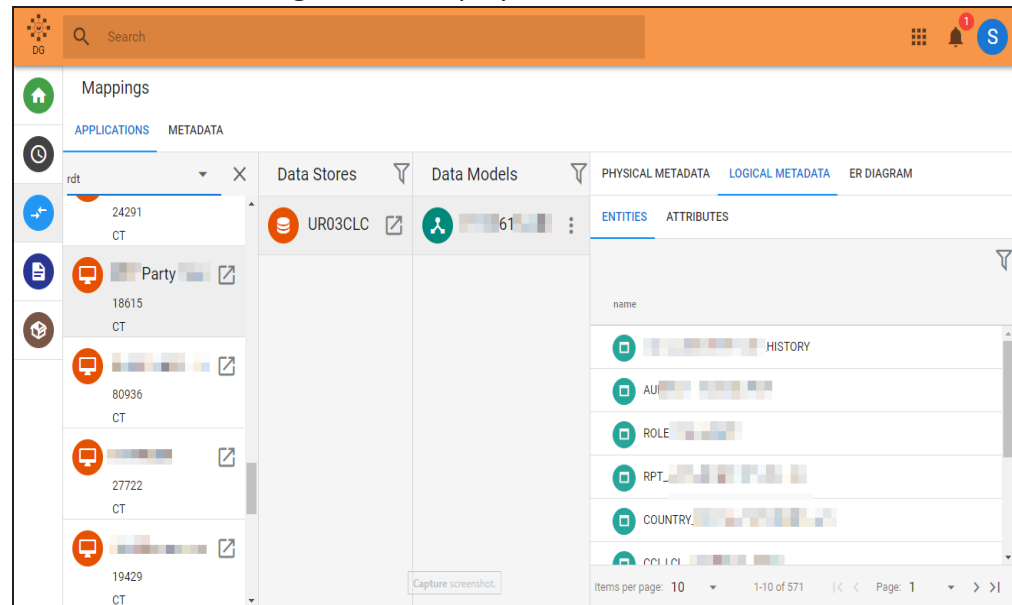
1. On the navigation bar on the left, click .

This opens the Mappings page.

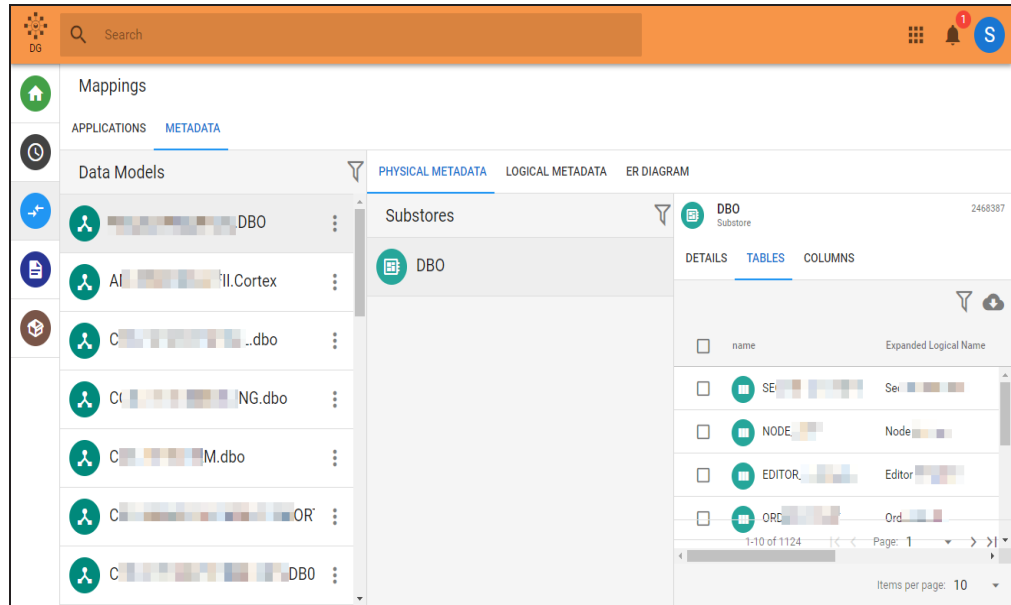


2. You can view mappings on either of the Applications or Metadata tabs.
  - a. The Applications tab enables you to navigate to tables and columns starting from an application. To view mappings using the Applications tab, do the following:

- i. Select an application.  
A list of associated data stores is displayed.
- ii. Select the data store for which you want to view the mapping.  
The associated data models and the corresponding physical and logical metadata, and ER diagram are displayed.



- iii. Click each tab to view corresponding metadata, such as tables, column, entities, attributes, and ER diagram.
  - iv. Select a table, column, entity or attribute to view further information.
- b. The Metadata tab enables you to navigate to tables and columns starting from a data model. To view mappings using the Metadata tab, do the following:
- i. Select a data model.  
The associated substores and the corresponding physical and logical metadata, and ER diagram are displayed on separate tabs.



- ii. Click each tab to view corresponding metadata, such as tables, column, entities, attributes, and ER diagram.
- iii. Select a table, column, entity or attribute to view further information.

## Viewing Reports


Reports are generated after a reverse engineering (RE) job is completed successfully.

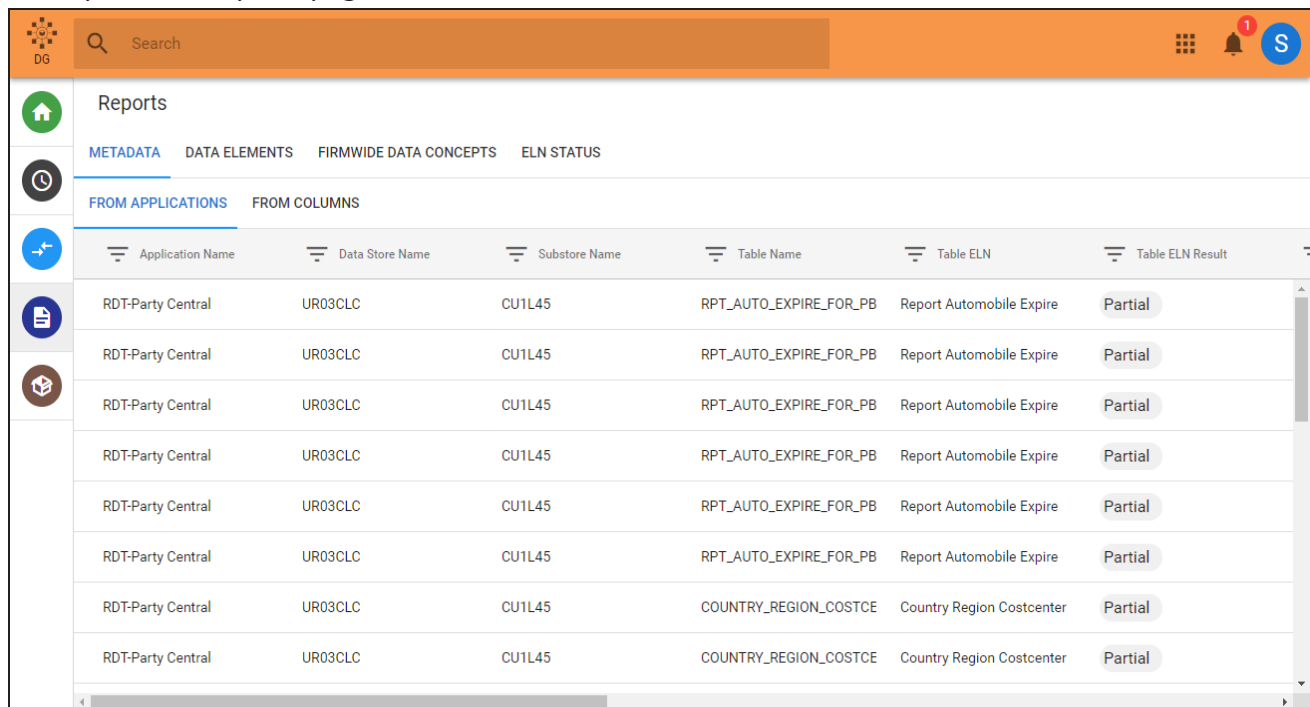
Using the Reports page, you can view the following:

- Relation of columns and tables to applications
- Mapping of columns to data elements
- Mapping of tables to firmwide data concepts
- Expanded Logical Name (ELN) process status
- ELN results

You can query the displayed information using the sorting and filtering features.

To view a report, do the following:

1. On the navigation bar on the left, click  .  
This opens the Reports page.



Application Name	Data Store Name	Substore Name	Table Name	Table ELN	Table ELN Result
RDT-Party Central	UR03CLC	CU1L45	RPT_AUTO_EXPIRE_FOR_PB	Report Automobile Expire	Partial
RDT-Party Central	UR03CLC	CU1L45	RPT_AUTO_EXPIRE_FOR_PB	Report Automobile Expire	Partial
RDT-Party Central	UR03CLC	CU1L45	RPT_AUTO_EXPIRE_FOR_PB	Report Automobile Expire	Partial
RDT-Party Central	UR03CLC	CU1L45	RPT_AUTO_EXPIRE_FOR_PB	Report Automobile Expire	Partial
RDT-Party Central	UR03CLC	CU1L45	RPT_AUTO_EXPIRE_FOR_PB	Report Automobile Expire	Partial
RDT-Party Central	UR03CLC	CU1L45	RPT_AUTO_EXPIRE_FOR_PB	Report Automobile Expire	Partial
RDT-Party Central	UR03CLC	CU1L45	COUNTRY_REGION_COSTCE	Country Region Costcenter	Partial
RDT-Party Central	UR03CLC	CU1L45	COUNTRY_REGION_COSTCE	Country Region Costcenter	Partial

2. Click each tab (Metadata, Data Elements, Firmwide Data Concepts, and ELN Status) to view the corresponding reports.
  - **Metadata:** The information displayed on this tab is an overview of metadata from applications to columns, and is a result of RE. It enables you to view metadata of an application, data store , or substore; from the application or columns perspective. Each perspective shows how data is related, the ELNs for tables or columns, additional column information, and the link to the data model on the Web Portal.
  - **Data Elements:** The information displayed on this tab is an overview of metadata from applications to data elements, and is a result of mappings you make between columns to data elements. You can view the information from the application or data element perspective. Each perspective shows how data is related, the ELNs for tables or columns, and additional column information.



- **Firmwide Data Concepts:** The information displayed on this tab is an overview of metadata from applications to firmwide data concepts, and is a result of mappings you make between tables to firmwide data concepts. You can view information from the application or firmwide data concept perspective. Each perspective shows how data is related and the ELNs for tables.
- **ELN Status:** This tab displays information about ELN generation and its results. you can view how many tables and columns have ELNs generated and how many do not. Under results, you can view whether a table or column ELN is "Not Possible", "Partial", or "Full".