

**Oracle® Product Data Quality**

PIM Connector User's Guide

Version 5.6.1

**E23405-01**

April 2011

Oracle Product Data Quality PIM Connector User's Guide, Version 5.6.1

E23405-01

Copyright © 2001, 2011 Oracle and/or its affiliates. All rights reserved.

Primary Author: Lorna Vallad

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this software or related documentation is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle USA, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

This software is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications which may create a risk of personal injury. If you use this software in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure the safe use of this software. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software in dangerous applications.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

This software and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

---

---

# Contents

Audience.....	v
Related Documents .....	v
Conventions .....	vi
<b>1 Overview</b>	
<b>2 Creating the Initial Semantic Model</b>	
1. Determining the Domain.....	2-1
2. Processing the Item Class Import from Excel .....	2-4
3. Using the AutoBuild Process in Excel.....	2-5
4. Reviewing Semantic Model Report.....	2-6
5. Reviewing Semantic Model .....	2-8
6. Refining the Semantic Model.....	2-8
7. Setting Up Standards and Match Rules.....	2-9
8. Testing a Data Lens on Sample Data and Create Base File.....	2-10
<b>3 Importing and Mapping an Alternate Catalog</b>	
1. Defining an Alternate Catalog.....	3-1
2. Processing Alternate Catalog Import from Excel .....	3-3
3. Using the AutoBuild Process in Excel.....	3-4
4. Reviewing Semantic Model Report.....	3-5
5. Reviewing Alternate Catalog.....	3-6
Creating Semantic Key 1 Cache.....	3-6
<b>4 Cleansing and Matching Batches</b>	
1. Creating a Batch.....	4-1
External Data Source.....	4-1
Production Pull Data Source .....	4-2
Processing the Interface Batch.....	4-2
2. Reviewing the Batch Processing Summaries.....	4-3
3. Reviewing Ready for Load .....	4-4
4. Reviewing the Alternate Catalog .....	4-5
5. Reviewing Duplicates Within the Batch.....	4-6
6. Reviewing Matches Against PIM.....	4-9
7. Reviewing Items for Enrichment .....	4-10

8. Reviewing Exceptions .....	4-10
9. Reviewing Style Matches .....	4-11
10. Reviewing the Release Batch .....	4-12
11. Reviewing Error Codes.....	4-12

## **A Installing the Client Software**

---

---

# Preface

This guide highlights the core process steps and features of Oracle Product Data Quality.

By extending the Oracle Product Information Management Data Hub with data quality capabilities, Oracle enables customers to improve the accuracy, completeness, and integrity of their master product data.

This guide uses the Vision Demo Database as a means to provide access to a common set of data, which is available to use as examples and validate the setup and configuration of the Oracle DataLens Server. It uses the Item Class Category Name of Capacitors under the Item Class of Passives.

To understand all of the features presented, you must use this guide in conjunction with the Oracle Product Data Quality documents listed in "[Related Documents](#)" on page -v.

You must have the Oracle Product Data Quality client software installed on your computer.

---

---

**Note:** Oracle Product Data Quality is only certified with the Oracle Site Hub and Oracle Product Hub for Retail Media Pack.

---

---

## Audience

You should have a basic understanding of the DataLens Technology.

This manual is designed for the new users of the Oracle Product Data Quality:

- Business Analysts
- Subject Matter Experts (SME)
- IT Administrators

## Related Documents

For more information, see the following documents in the documentation set:

- The *Oracle Product Data Quality PIM Connector Implementation Guide* provides installation and configuration of Oracle Product Data Quality Connector.
- The *Oracle Product Data Quality Oracle DataLens Server Installation Guide* provides detailed Oracle Product Data Quality Oracle DataLens Server installation instructions.

- The *Oracle Product Data Quality Oracle DataLens Server Administration Guide* provides information about installing and managing an Oracle DataLens Server.
- The *Oracle Product Data Quality COM Interface Guide* provides information about installing and using the Oracle DataLens Server COM APIs.
- The *Oracle Product Data Quality Java Interface Guide* provides information about installing and using the Oracle DataLens Server Java APIs.
- The *Oracle Product Data Quality Application Studio Reference Guide* provides information about creating and maintaining Data Service Applications (DSAs).
- The *Oracle Product Data Quality AutoBuild Reference Guide* provides information about creating initial data lens based on existing product information and data lens knowledge.
- The *Oracle Product Data Quality Knowledge Studio Reference Guide* provides information about creating and maintaining data lenses.
- The *Oracle Product Data Quality Governance Studio Reference Guide* provides information about creating and maintaining Data Service Applications (DSAs).
- The *Oracle Product Data Quality Glossary* provides definitions to commonly used Oracle Product Data Quality technology terms.
- The *Oracle Product Data Quality Services for Excel Reference Guide* provides information about creating a DSA based on data contained in a Microsoft Excel spreadsheet.
- The *Oracle Product Data Quality Task Manager Reference Guide* provides information about managing tasks created with the Task Manager or Governance Studio applications.

See the latest version of this and all documents listed at the Oracle Product Data Quality Documentation Web site at:

[http://download.oracle.com/docs/cd/E20593\\_01/index.htm](http://download.oracle.com/docs/cd/E20593_01/index.htm)

## Conventions

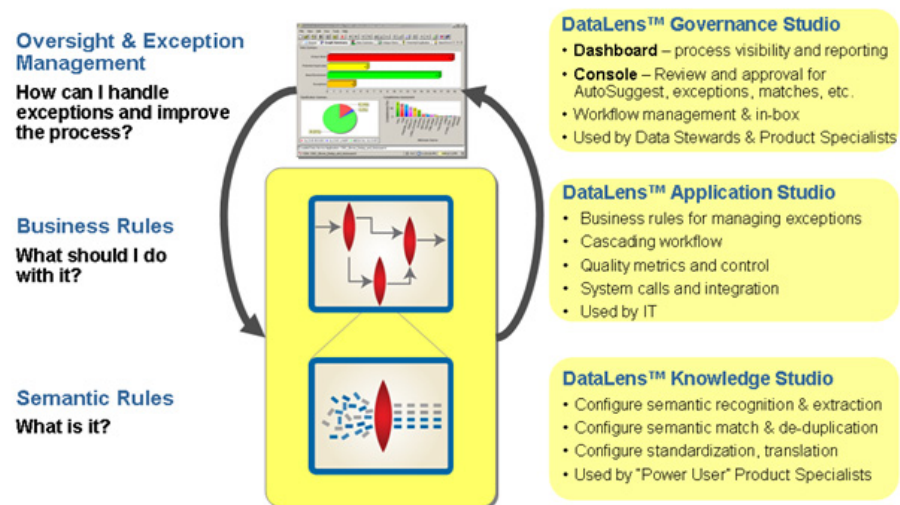
The following text conventions are used in this document:

Convention	Meaning
<b>boldface</b>	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
<i>italic</i>	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
monospace	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, text that you enter, or a file, directory, or path name.
<b>monospace</b>	Boldface, monospace type indicates commands or text that you enter.

## Overview

Oracle DataLens Server is built on industry-leading DataLens™ Technology to standardize, match, enrich, and correct product data from different sources and systems. The core DataLens Technology uses patented semantic technology designed from the ground up to tackle the extreme variability typical of product data.

Oracle Product Data Quality uses three core DataLens Technology modules: Governance Studio, Knowledge Studio, and Application Studio. The following figure illustrates the process flow of these modules.



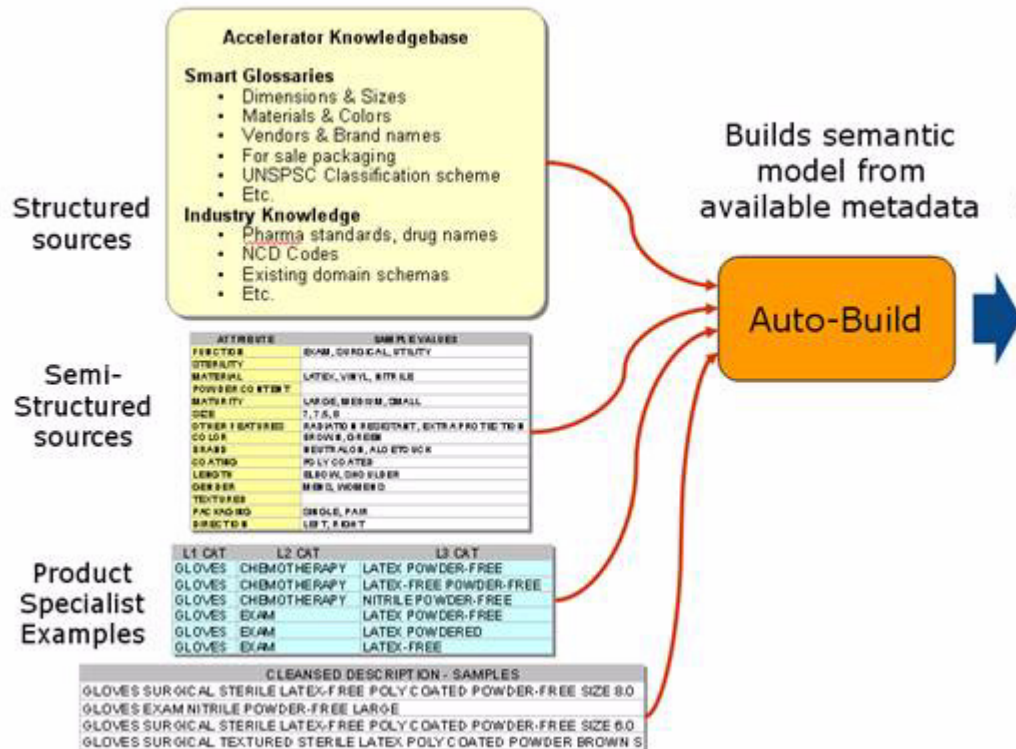
The process uses Oracle Product Data Quality AutoBuild, which is a component of the DataLens Technology to quickly leverage the existing Product Data Hub set of information and the Oracle Product Data Quality Smart Glossaries to create an initial Semantic Model (data lens) specific to your enterprise content.

Oracle Product Data Quality ships with a set of application components to allow you to test and ensure proper system configuration. The default data lens to process items is named 001\_Capacitors\_Demo\_V2. This data lens contains semantic models that work with the Capacitors Item Class category in the Vision Database installed with most Oracle Apps Unlimited (AU) Product Information Management (PIM) systems. The remainder of the document uses these samples shipped with the product.

AutoBuild constructs a data lens by examining the available product data examples. Given sufficient information, AutoBuild can accomplish the following.

1. Construct a full Item Definition hierarchy, complete with required, scoring, and optional attributes
2. Construct rich term and phrase recognition rules
3. Provide an initial set of classification rules

AutoBuild offers a familiar, easy-to-use graphical wizard interface that steps you through the process from start to finish.



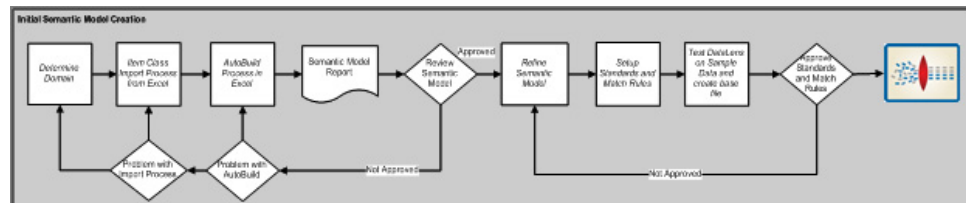


## Creating the Initial Semantic Model

You create an initial Semantic Model using AutoBuild, which is extracts the metadata from the Oracle Item Class tables and samples from the Product Data Hub Production tables. The process takes a single input parameter for Item Class Category Name and uses another parameter called Display Type Format. The Display Type Format is contained in the Profile Options value called "SemanticModel". For more information about configuring this option, see *Oracle Product Data Quality PIM Connector Implementation Guide*.

The Item Class category will determine which sections of the Item Class Hierarchy will be extracted and the Display Format will determine which attribute metadata to extract to create the initial Semantic Model.

The Product Data Hub contains both information about the Item Classification and Attributes (Metadata) as well as sample data. The AutoLearn process leverages this information to create the initial semantic model. A Semantic Model is equivalent to a fully configured Data Lens with the Item Definition, attributes, phrases and terms as well as the standardizations and match rules. The following figure illustrates the process.



### 1. Determining the Domain

You must determine what domain (Item Class) they would like to use to create the Semantic Model. This requires the setup of an Oracle Item Class Category Name.

#### Oracle Item Class Category Name Setup

The Semantic Model creation process relies on each Item Class Category Name being setup correctly and named correctly. The Semantic Model is associated with the Item Catalog hierarchy. The following shows the "Capacitor" hierarchy, as laid out in the Vision Operations Item Catalog.

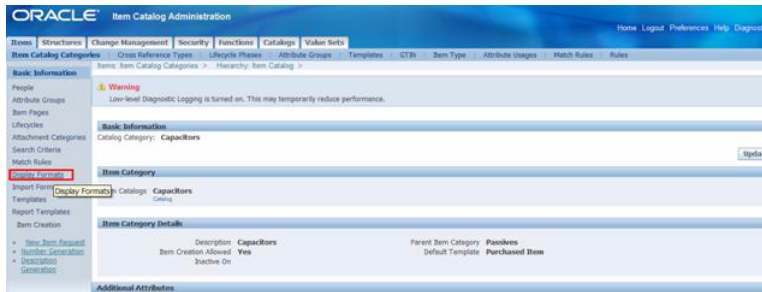
- PLM High Tech
  - Electronic Components
    - Passives
      - Capacitors

Click **Setup Workbench**. Search the Item Catalog Category for the category you want to create a semantic model for by entering a search term then click **Go**. In this example, search for "Capacitors".



Click **Capacitors**.

From the left navigation menu, click **Display Formats** to create the Semantic Model definition created from a set of attributes available in the Product Data Hub.



The Display Formats page allows you to create, copy, edit (update), or delete a Display Model.



Create a Display Format as follows:

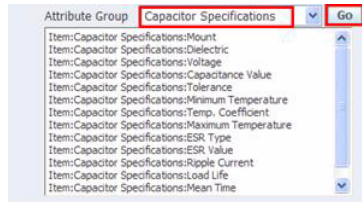
1. Click **Create**.
2. Enter **SemanticModel** as the name.
3. Enter **DataLens Semantic Model Attributes** for the description.
4. Select **100** for the Number of Rows Displayed.
5. Select the **Set as Default** check box.



6. Create the Display Format by selecting **Attributes** from the Attribute Columns list on the left side (there are a set from the Primary Attribute Group by default in the right list).



- a. From the **Business Entity** list, select **Item**.
- b. Each Attribute Group has a set of Attributes values, for example by choosing **Capacitor Specifications** and clicking **Go** displays the set of attributes associated with that Attribute Group.

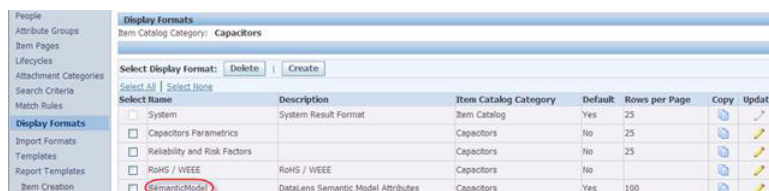


Highlight an individual attribute and click the **Move** button to move them to the **Selected Columns** list.

7. Select all of the **Item:Capacitor Specifications** attributes using **Move All** button. The attributes should include the ones as shown in the following table:

Business Entity	Attribute Group	Attribute
Item	Capacitor Specifications	Capacitance
Item	Capacitor Specifications	Dielectric
Item	Capacitor Specifications	ESR_Type
Item	Capacitor Specifications	ESR_Value
Item	Capacitor Specifications	Load_Life
Item	Capacitor Specifications	Maximum_Temperature
Item	Capacitor Specifications	Mean_Time
Item	Capacitor Specifications	Min_Temperature
Item	Capacitor Specifications	Mount
Item	Capacitor Specifications	Ripple_Current
Item	Capacitor Specifications	Temp_Coefficient
Item	Capacitor Specifications	Tolerance1
Item	Capacitor Specifications	Voltage

8. Click **Apply**.



A new Display Format named "SemanticModel" with all of the necessary attributes is created for input into the AutoBuild process. The set of attributes should correlate to the spreadsheet output tab 10\_AutoBuild in the CreateSemanticModels.xls spreadsheet.

---



---

**Note:** The AutoBuild spreadsheet tab has attribute name / value pairs where the attributes names are in the attr\_name column and they run vertically.

---



---

## 2. Processing the Item Class Import from Excel

You use the CreateSemanticModels.xls Excel spreadsheet template to import the Item Class (Category) and Attribute Metadata from the Oracle Product Information Management Data Hub system to create the import files for AutoBuild.

---



---

**Note:** The Template relies on the Oracle Product Data Quality Services for Excel.

---



---

1. Open the CreateSemanticModels.xls spreadsheet template to the Sheet 1 tab.
2. From the Services for Excel Add-In tab, click the **Submit Job** button to launch the ORA\_AU\_Autobuilder DSA to create the AutoBuild input files for specific categories for use by AU PIM.
3. Enter the Item Class Category parameter (for example, "Capacitors") associated with the Semantic Model that you are building and click **OK**.

The job is submitted to the server for processing.

4. Click the **Get Job Results** button to retrieve the results from the completed job. When the job is retrieved, three additional tabs are created, 10\_Autobuild and 20\_SampleData.

- The 10\_Autobuild tab contains the following columns:

Columns	Definition
parent_catalog_group_id	System ID
parent_catalog_group	Item Class Category Parent Hierarchy Node Name
catalog_group_id	System ID
catalog_group	Item Class Category Name
attr_id	System ID
attr_name	Attribute name
attr_value	Attribute value

- The 20\_SampleData tab contains the following columns

Columns	Definition
catalog_group	Item Class Category Name

Columns	Definition
sku	Part number
description	Attribute description

The 10\_AutoBuild tab should have at least one of the attributes (it may be with no data) listed from the Display Format in the Semantic Model created in the first step.

### 3. Using the AutoBuild Process in Excel

AutoBuild constructs a data lens from the available product data metadata and example data (the server job pulls a distinct set of data) shown in the 10\_AutoBuild tab. Given the correctly formatted information, AutoBuild will accomplish the following.

- Construct a full Item Definition hierarchy based on the Item Catalog Category, complete with required, scoring, and optional attributes
- Construct rich term and phrase recognition rules
- Provide an initial set of classification rules
- Create Match Rules (Exact and Close Match)
- Create Standardizations Types

From the 10\_AutoBuild tab, start the AutoBuild process by clicking the **AutoBuild** button on the Services for Excel toolbar.

AutoBuild provides an easy-to-use graphical wizard interface to a four step process to build the data lens.

Enter the following default settings into the AutoBuild wizard as follows:

Step	Value
<b>Step 1: Identify Category Information</b>	
Generate a new DataLens	Selected
Smart Glossary	DLS_Import_Template
List of category columns	1-4
<b>Step 2: Identify Attribute Information</b>	
Required Attributes	
Add Required Attributes	Selected
Item Name Column(s)	4
Scoring Attributes	
Add Scoring Attributes	Selected
Attribute names are in the same row at the category	Selected
Row contains Attribute Name/Value pairs	Selected
Scoring Attributes	6,7
Single term columns	6,7

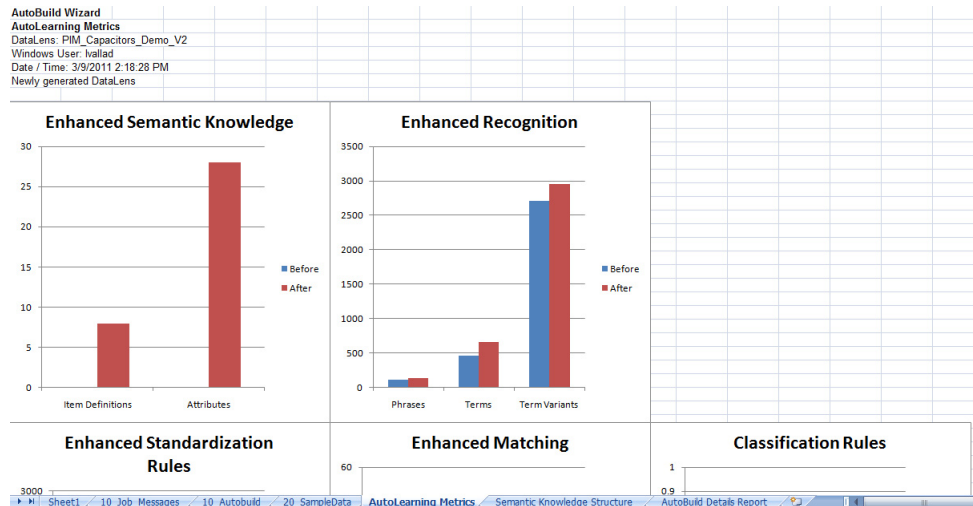
<b>Step</b>	<b>Value</b>
Standardizations	Select <b>Attribute Details</b> <<
Create Replace All Rules	Selected
<b>Step 3: Identify Sample Data Information</b>	
Sample data file	Selected
Use current file	Selected
Sample Data Worksheet	20_SampleData
Sample Data Rows and Columns	
Id Column	2
Group By Column	1
List of Sample Data Columns	3
<b>Step 4: Identify Sample Data Information</b>	
Outputs	
Data Lens Options	
Generate / Update Data Lens	Selected
Open New / Updated Data Lens	Selected
New / Updated Data Lens	<i>Data Lens Name</i>
Output Details	Select <b>Output Details</b> <<
Create Match Rules	Selected
Define Exact Match	
Define Exact Match	4
Define Close Match	1
Generate Report	Selected

The initial data lens will be built with the Item Definition (Category) and attribute structure shown in the pre-configured AutoBuild settings. A set of reports added to the `CreateSemanticModels.xls` spreadsheet as tabs. These reports include, `AutoLearning Metrics Report`, `Semantic Knowledge Structure`, and `AutoBuild Details Report`.

## 4. Reviewing Semantic Model Report

Review the following AutoBuild reports to validate the data lens creation:

The `AutoLearning Metrics` report tab provides two sets of graphs that provide a before and after view of the Semantic Knowledge and Text Recognition. The initial data lens uses the `DLS_Import_Template` Smart Glossary as the starting point (before).



The Semantic Knowledge Structure tab defines the Item Definition and the associated Required / Scoring Attributes, terms and data. The report has three sections: Attribute Phrases and Terms. You should validate that each of them correlate with the 10\_AutoBuild tab.

AutoBuild Wizard  
 Semantic Knowledge Structure  
 DataLens: PIM\_Capacitors\_Demo\_V2  
 Windows User: wallad  
 Date / Time: 3/9/2011 2:18:28 PM  
 Newly generated DataLens

[Item Definitions](#)  
[Item Definitions with Attributes](#)  
[Phrases](#)  
[Terms](#)  
[Alternate Catalog](#)

Item Definitions			
Level: 1	Level: 2	Level: 3	
Item_571	Item_571_Capacitors	Item_572	Capacitor_Networks

[New Meta-Data](#)  
[Renamed Meta-Data](#)  
[Existing Meta-Data](#)  
[Not Present in Meta-Data](#)

[Item Definitions with Attributes](#)

[New Meta-Data](#)  
[Renamed Meta-Data](#)  
[Existing Meta-Data](#)  
[Not Present in Meta-Data](#)

The AutoBuild Details Report tab is a high level summary report showing the knowledge creation results.

AutoBuild Wizard  
 AutoBuild Details Report  
 DataLens: PIM\_Capacitors\_Demo\_V2  
 Windows User: halled  
 Date / Time: 3/9/2011 2:18:29 PM  
 Newly generated DataLens

[Totals](#)  
[Performance](#)  
[Lens Refinements](#)  
[Item Definition Refinements](#)  
[Phrase Refinements](#)  
[Term Refinements](#)  
[Classification Refinements](#)  
[AutoBuild Warnings](#)

Totals	
<b>Input Meta-Data Metrics</b>	
Unique Categories found	8
Leaf Categories found	2
Attributes found	28
Attribute values found	734
Category Depth	4
<b>AutoLearning Metrics</b>	
New Item Definitions	8
New Attributes	28

Sheet1 | 10\_Job\_Messages | 10\_AutoBuild | 20\_SampleData | AutoLearning Metrics | Semantic Knowledge Structure | AutoBuild Details Report

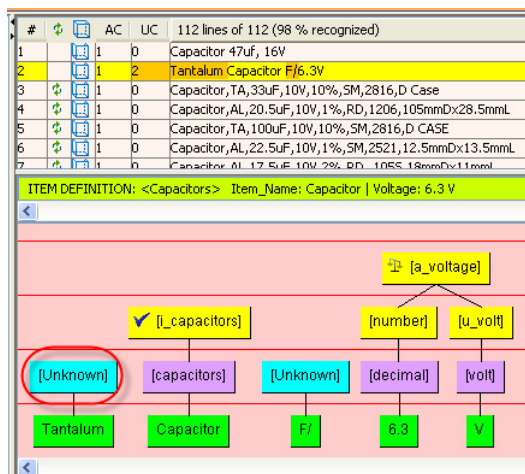
These reports should correctly represent the results of the AutoBuild knowledge building process and should correspond to the data lens knowledge. For more information about AutoBuild, see *Oracle Product Data Quality AutoBuild Reference Guide*.

## 5. Reviewing Semantic Model

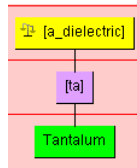
Review the data lens to confirm proper creation. The data lens should have the correct Item Definition, Attributes Phrases, and Terminology as defined by the AutoBuild process.

## 6. Refining the Semantic Model

You should review the data lens to define any undefined terminology to complete the process. The following diagram shows an example of the term "Tantalum" not being recognized and shown as "Unknown". A SME would know that this is a variant of 'ta', which is a type of dielectric. You would make the addition to refine the data lens.



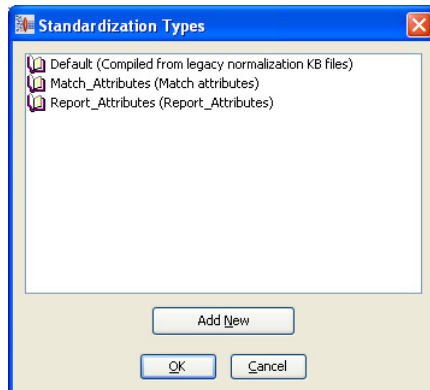




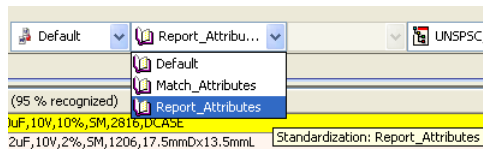
For more information, see *Oracle Product Data Quality Knowledge Studio Reference Guide*

## 7. Setting Up Standards and Match Rules

Setup and review the data lens standardizations. There should be three Standardizations: Default, Match\_Attributes, and Report\_Attributes. You can access them from the **DataLens** menu by selecting the **Standardization Types**.



You can also view the available standardizations from the **Standardizations** list on the toolbar.

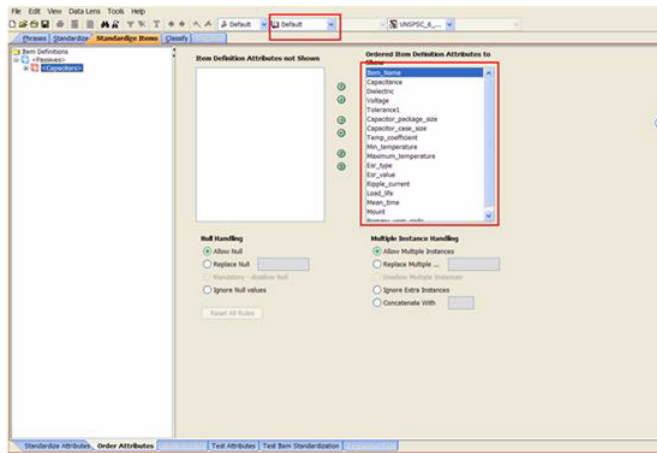



---

**Note:** The attribute order is the same order as the name/value attribute order in the AutoBuild tab, which is based on the Display Format order. For the Match\_Attributes standardization, there should be two Match weights created (Exact\_Match with four attributes and Close\_Match with one attribute).

---

For example, default Standardization would have the following attribute order:

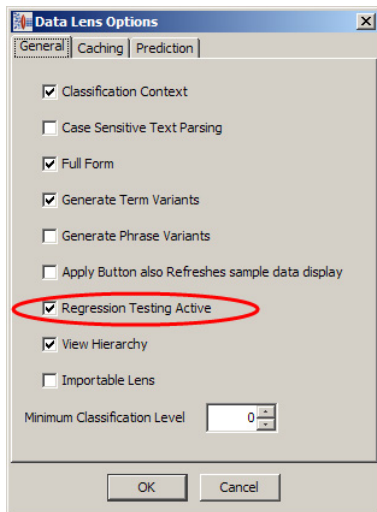


You should review and change the attribute ordering, create phrase standardizations and replacements appropriate to the data lens.

## 8. Testing a Data Lens on Sample Data and Create Base File

The purpose of regression testing is to validate that any maintenance to the data lens has not created any ambiguities or issues compared to the phrases and terms that were previously defined in a prior version of the data lens. This important step should be performed after changes have been made to a mature data lens.

Regression testing can be accessed from the phrases, standardization and classification tabs. You should set the **Regression Testing Active** option from the **DataLens Options** to create the base file to enable Regression testing.

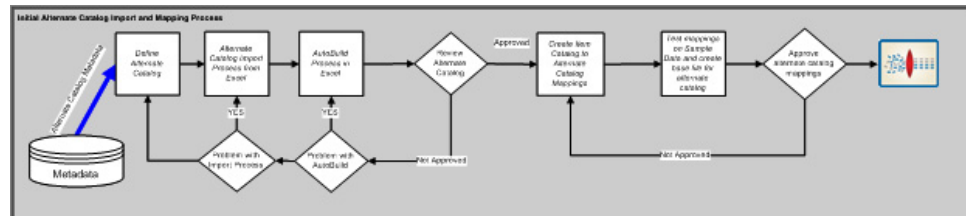


For more information about setting up and using Regression tests, see *Oracle Product Data Quality Knowledge Studio Reference Guide*.

Once you are happy with the data lens works, you should contact your Administrator to associate this new data lens with the cleansing and matching process.

## Importing and Mapping an Alternate Catalog

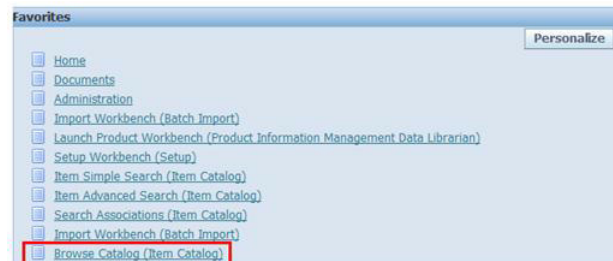
You create an alternate catalog mapping through the AutoBuild program, which extracts the Oracle Alternate Catalog Hierarchy structures metadata from the Product Data Hub Production tables. The process takes a single input parameter for Item Class Category Name. The Item Class category will determine which sections of the Item Class Hierarchy will be extracted to create an export that can be used by AutoBuild to create a schema file.



### 1. Defining an Alternate Catalog

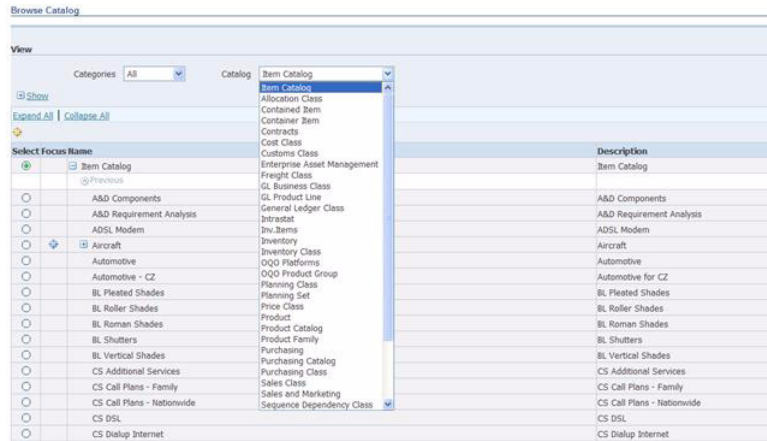
You must determine what domain (other catalog) they would like to associate with the previously created Semantic Model. This requires you to know what catalogs exist in the Oracle Product Data Hub in order to select the one to associate with the existing Oracle Item Class Category Name.

Select **Browse Catalog**.



The Oracle Product Data Hub has a set of preconfigured catalogs. From the home page, you can browse the catalogs and view the active catalogs from the **Catalog** list.

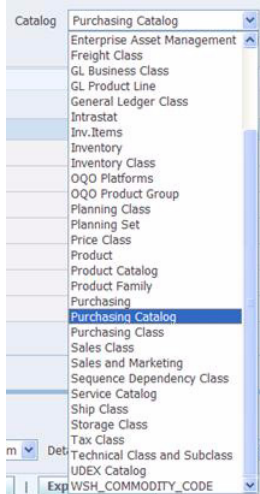
## 1. Defining an Alternate Catalog



The following shows the Capacitor hierarchy as laid out in the Vision Operations Item Catalog.

- Item Catalog
  - PLM High Tech
    - Electronic Components
      - Passives
        - Capacitors

The default configuration includes the Purchasing Catalog as an alternate catalog for the Capacitors Semantic Model. Select **Purchasing Catalog** from the **Catalog** list.



The Purchasing Catalog has a slightly different hierarchy with one less level.

- Purchasing Catalog
  - Electronic Components
    - Passives
      - Capacitors

Purchasing Catalog	Purchasing Catalog
Electromechanical Components	Electromechanical Components
Electronic Components	Electronic Components
Discretes	Discretes
IC's	IC's
Optoelectronics	Optoelectronics
Passives	Passives
Capacitors	Capacitors
Circuit Protection	Circuit Protection
Encoders	Encoders
Frequency Controllers	Frequency Controllers
Inductors	Inductors
RC Networks	RC Networks
Resistors	Resistors
Signal Transformers	Signal Transformers

An alternative catalog is a new user defined classification to add to the Semantic model (data lens) to allow the association of individual item with more than one catalog.

## 2. Processing Alternate Catalog Import from Excel

You use the `ImportAlternateCatalogs.xls` Excel spreadsheet to import the Alternate Category from the Oracle Product Data Hub system to create the import files for AutoBuild.

1. Open the `ImportAlternateCatalogs.xls` Excel spreadsheet template to Sheet 1.
2. From the Services for Excel Add-In tab, click the **Submit Job** button create the AutoBuild input files for specific alternate category for use by AU PIM.
3. Enter the Alternate Catalog name parameter. In this example, enter **Purchasing Catalog**, and the click **OK**. The job will be submitted to the server for processing.

You can use any Catalog that is listed in Oracle E-Business, but there should a logical relationship for the selection.

- i. Purchasing Catalog
  - ii. Product Catalog
  - iii. So on...
4. Click the **Get Job Results** button to retrieve the results from the completed job.

When the job is retrieved, it creates one additional tab, `001_Alternate_Catalog`, which contains the following columns.

Columns	Definition
catalog	Alternate Catalog name
parent_catalog_id	System ID
parent_category_desc	Alternate Catalog hierarchy descriptions
category_id	System ID
catetory_name	Alternate Catalog hierarchy category names

The output tab, `001_Alternate_Catalog`, will be input to the AutoBuild process.

### 3. Using the AutoBuild Process in Excel

AutoBuild constructs a new user defined classification into a data lens from the available category information shown in the 001\_Alternate\_Catalog tab. Given the correctly formatted information, AutoBuild will construct a new user defined Classification with the same name as the catalog name (for example, "Purchasing\_Catalog") using the name in the catalog column. You should start the AutoBuild process from the 001\_Alternate\_Catalog tab.

The "catalog" column name is a special name for the AutoBuild process and must be used to construct the classification (Alternate Catalog).

1	catalog	parent_category_id	parent_category_desc	category_id	category_name
---	---------	--------------------	----------------------	-------------	---------------

AutoBuild uses the numeric structure and relationship of the parent\_category\_id and category\_id to define the classification hierarchy.

1	catalog	parent_category_id	parent_category_desc	category_id	category_name
---	---------	--------------------	----------------------	-------------	---------------

For the purposes of this flow, the Administrator must have included the data lens in the process.

**Note:** The process can automatically handle the creation of three alternate catalogs. If you require more than three alternate catalogs, then contact your Oracle DataLens Server Administrator.

The default AutoBuild settings for each step are as follows:

Step	Value
<b>Step 1: Identify Category Information</b>	
Smart Glossary	Use an existing data lens aside from Smart Glossaries (For example, DLS_ prefix)
List of category columns	3,5
Category column layouts	
Multi-Column List of Category Code/Title pairs	Selected
<b>Step 2: Identify Attribute Information</b>	
Nothing Entered	
Required Attributes	
Add Required Attributes	
Item Name Column(s)	
Scoring Attributes	
Add Scoring Attributes	
Attribute names in same row	
Row contains Attribute Name/Value pairs	
Scoring Attributes	
Single term columns	
<b>Step 3: Identify Sample Data Information</b>	
Nothing Entered	

Step	Value
Select Sample Data Source	
Use DataLens Knowledge Studio to collect samples	Selected
<b>Step 4: Identify Sample Data Information</b>	
Outputs	
Data Lens Options	
Generate / Update Data Lens	Selected
Open New / Updated Data Lens	Selected
New / Updated Data Lens	<i>Data Lens Name</i>
Output Details	Select <b>Output Details</b> <<
Create Classification Type	Selected (Dimmed)
Type Name	Catalog name (shown in catalog column)
Create Match Rules	Dimmed
Define Exact Match	
Exact_Match	4 (Dimmed)
Close_Match	1 (Dimmed)
Generate Report	Selected

---

**Note:** The data lens can be anything to test the creation of you defined classification, but must be correct data lens in order to affect the entire process as previously described.

---

When the process completes, the selected data lens will have a new user-defined classification. AutoBuild will also create a set of reports added to the spreadsheet as tabs. These reports include, *AutoLearning Metrics Report*, *Semantic Knowledge Structure*, and *AutoBuild Details Report*.

## 4. Reviewing Semantic Model Report

Review the AutoBuild reports to validate the data lens creation.

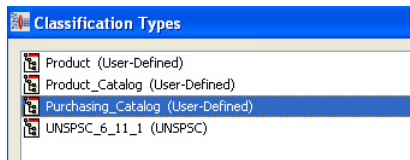
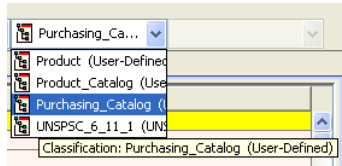
AutoBuild Meta Data Report defines the Item Definition and the associated Required / Scoring Attributes, terms and data. There are three sections (Attribute, Phrases and Terms) to the report. You should validate that each of them correlate with the data in the *AutoBuild* tab.

AutoBuild Summary Report provides two sets of graphs that provide a before and after view of the Semantic Knowledge and Text Recognition. The initial lens may be the *DLS\_Import\_Template* Smart Glossary or the target data lens as the starting point (before).

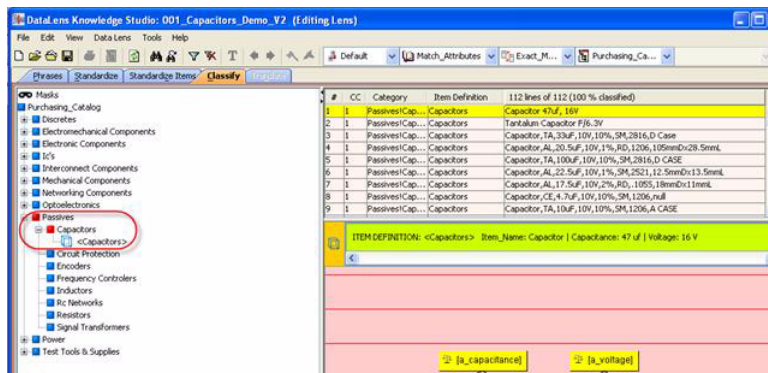
AutoBuild Details Report is a high level summary report showing the knowledge creation results. These reports should correctly represent the results of the AutoBuild knowledge building process and should correspond to the data lens knowledge.

## 5. Reviewing Alternate Catalog

You should review the data lens to confirm that your defined classification is correct. The first step is to determine if the alternate catalog is present in the data lens. The new alternate catalog will appear as a **(User-Defined)** type (for example, "Purchasing\_Catalog"). You can access it from the **DataLens** menu, by selecting **Classification Types** or by clicking the list from the toolbar:



The final step is to open up the **Classify** tab and relate the Semantic Model (Item Definition) to the Category. In this case, Capacitors is related to Capacitors to create an alternate catalog assignment. For more information about setting up and using Classification Types, see *Oracle Product Data Quality Knowledge Studio Reference Guide*.



## Creating Semantic Key 1 Cache

Oracle Product Data Quality uses patented Semantic technology to create a set of standardized attributes that are output to generate a set of master data or Semantic Key 1 Cache. The cache represents a "semantic" index created from the extracted and standardized attribute information from the PIM.

The semantic cache is used to rapidly match (compare) requested products to the items found in the cache. Any matches are returned to enable further processing such as form, fit or functional equivalents or duplicate detection to protect the PIM from the addition of identical records.

Open the Governance Studio and run the `Update_Semantic_Cache` job. It will ask for a Category as input. Enter the Category (for example, Capacitors). Once this job has run, it produces the Semantic Key 1 Cache for that Category with typically no results shown in a Governance Studio tab for successful items. Two review tabs, `Unassigned_Semantic_Key` and `Unassigned_Item_Class`, are created to provide you with information about exceptions.



At the end of the process, you will have a complete Semantic Key 1 Cache. This is really an IT function and can only be confirmed by looking at the database. For more details, see *Oracle Product Data Quality PIM Connector Implementation Guide*.

The process is typically run on a periodic basis (for example, nightly) to ensure that new updates to the PIM are included in the cache.



---

---

## Cleansing and Matching Batches

A newly created batch will be a status of "Active". When the batch is in process, the system will automatically change the batch status to "Pending" to ensure that the batch cannot be changed while it is in process. This is a critical step to ensure that the job can be correctly processed. The Governance Studio project will not produce any results and will generate an error on the **Error Codes** tab if the status is not set to "Active".

### 1. Creating a Batch

There are two types of batches that can be created for processing. The first type of batch is from an external source (outside the PIM). The second type of batch is a pull from production (within the PIM).

#### External Data Source

You can create a batch to process data that is external to PIM to load it into the following PIM interface tables:

- EGO\_ITM\_USR\_ATTR\_INTRFC
- MTL\_SYSTEM\_ITEMS\_INTERFACE
- MTL\_ALT\_CATEGORIES\_INTERFACE

The external data must be in an Excel spreadsheet containing an item id column and a description column. These columns are the defaults necessary to the AUPIM\_PROCESS\_EXTERNAL\_DATA DSA that is used for processing the data.

1. Open the Excel spreadsheet containing your external data.
2. From the Services for Excel toolbar, click the **Process Records** button.

If you have not logged into your Oracle DataLens Server, the **Oracle Product Data Quality Login** dialog box appears. Login to the server displayed or use the Change Server button to login to a different server.

When you have successfully logged into the server, your data is processed, inserted in the PIM interface table, and the batch is created within PIM. A status dialog box is displayed that identifies the JobID.

3. Click **OK**.
4. Review the 10\_Job\_Messages tab to verify that no processing errors occurred. If errors occur, resolve the errors in the DSA or data, and then reprocess the records.
5. In E-Business Suite, click **Import Workbench**.
6. Click **Go**.

All batches are displayed.

7. Review the list of batches and confirm that your external batch job is displayed. The JobID displayed in Excel is included in the Batch Name.
8. (Optional) Select the **Update** pencil (second to last column) to review the batch contents.

## Production Pull Data Source

The `Create_PIMDH_Production_Batch` Governance Studio project creates a new Batch, pulling data from the Production tables in the PIM.

1. Enter the name of the Item Category (for example, Capacitors).  
When the job finishes, it creates a new batch with data from the selected Item Category.
2. Review the **Batch Creation Results** tab to see the results of the job run, and note or copy the new Batch Id for later use.
3. Review the newly created Batch in Oracle Batch Interface table. The new batch uses the following defaults:

Field Name	Value
Batch Id	<i>New Batch Number (System Generated)</i>
Batch Name	DLS Production pull batch
Batch Type	Item
Source System	Product Information Management Data Hub
Assigned To	Albert Bakker
Batch Status	Completed

---

---

**Note:** The batch is set to the "Active" status in order to allow the system to correctly process the batch; otherwise, you an error is generated.

---

---

The results will be an Oracle batch with a set of imported rows in the **Confirmed** tab created from the production PIM data for the selected item category. They are considered Confirmed because they are the PIM.

## Processing the Interface Batch

This section describes the capabilities of each of the Process Interface Batch tabs.

You use the Batch Id to create a unique Governance Studio job to process the data pulled for the selected Category.

1. Open the `DataLens_Cleansing_and_Matching` project or a previously created project (for example, `DataLens_Cleansing_and_Matching-10003377`).
2. From the **File** menu, select **New Project From...**

You will be prompted to enter a Unique Id (for example, Batch Number plus a version number).

3. Enter the previously created Batch number (for example, 10003377\_V1) from the previously created Batch number as input to the Governance Studio job, click **Run**.

---

**Note:** The difference between a creating a job with data pulled from the PIM and data from an external source is that there is not a concept of checking for duplicates so those tabs are always be blank.

---

Once the `DataLens_Cleansing_and_Matching` Governance Studio job runs, it produces a series of output tabs that provide you with a way to process the batch using a series of processing tabs. An individual item uses the following process logic.

1. Error Codes	If the batch is not correctly prepared or not in the correct state to be processed.
2. Release Batch	Provides a mechanism to change the Oracle batch from "Pending" to "Active", allowing a user to review the completed changes.
3. Exceptions	If the record falls below a Quality Index of 20 (nothing recognized) then it becomes an exception.
4. Dups Within Batch	The process checks for duplicates within the original source batch and presents them to you to determine what should be done.
5. Match on Mfg Part	The process checks for duplicates by comparing the source manufacture name and part number against an existing record in the PIM.
6. Match Against PIM	The process checks for the source with the records in the PIM using the Semantic index cache to determine if the source is a duplicate to an existing record in the PIM.
7. Items for Enrichment	If a source record is not a duplicate and does not meet the Quality Index threshold then present it to you to enrich the records so that they may be of the necessary completeness to load into the PIM.
8. Ready for Load	Fully complaint, standardized and enriched records ready to be imported into the PIM
9. Alternate Catalog Review	All fully standardized records (meets the Quality Index threshold) may have an alternate catalog (classification) assigned based on the Semantic Model definition.
10. Style Matches	A final check against the PIM to determine if this item could be associated with an existing Style in PIM.

---

**Note:** Once the processing is complete, you will have to run Release Batch process from within Governance Studio job to reset the batch status from "Pending" to "Active" in order to review and confirm the changes made during this process.

---

## 2. Reviewing the Batch Processing Summaries

The Governance Studio **Graph Summary** tab provides a graphical summary of the state of the batch processing.

The **Data Summary** chart shows the totals by processing category. As you process items, the numbers will change and you can monitor this pane to verify the result. For

example, if you enrich a record to the point where it can pass the quality threshold for import into the PIM, it will move from the Items for Enrichment category to the Confirmed New Review category.

The **Classification Summary** chart displays the total in a pie chart by Classification. If there is only one type then only one type will show.

The **Completeness Assessment** charts displays information about the attributes. It shows the coverage of the attributes across the record set. If a particular attribute has 100% coverage then that individual attribute count (completeness) will equal the record set total.

### 3. Reviewing Ready for Load

The Governance Studio **Ready for Load** tab provides a view of the fully complaint, standardized and enriched records that met the quality requirements and are "import ready" and will be moved into the PIM through the interface table (batch) process.

Once you have reviewed and handled all of the items on the other tabs, bringing the totals to zero, the remainder will appear on the **Confirmed New Review** tab. You select one or more rows to complete the process, which updates the Interface Batch with this set of standardized high quality data, using the Batch Id and the ITUI (Interface Table Unique Id) as the key as follows:

1. In the Governance Studio, select the **Ready for Load** tab.
2. Select on from one to all of the rows in the tab.

	interface_table_unique_id	organization_id	source_syste
✓	67643	204	Agile Electroni
✓	67644	204	Agile Electroni
✓	67647	204	Agile Electroni
✓	67656	204	Agile Electroni
✓	67663	204	Agile Electroni
✓	67665	204	Agile Electroni
✓	67670	204	Agile Electroni

3. Click the **Apply Checked Rows** button to process the selected rows.

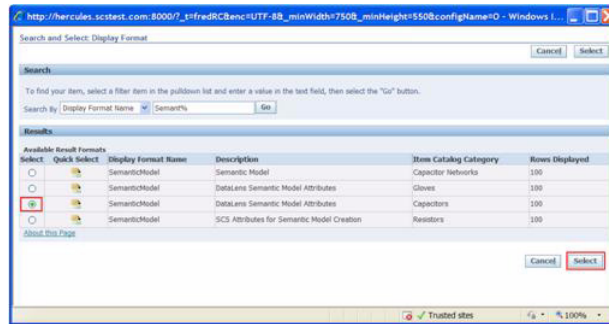
The selected rows turn blue indicating that the rows have been processed and that the rows are now unavailable for further processing.

4. Once the processing is complete, run the Release Batch process to reset the batch status from "Pending" to "Active" then review and confirm the changes made during this process. The job uses the batch id as input.
5. Confirm that the data is set correctly for the selected row(s) by going back into batch in the Oracle Item Import Workbench (assumes batch set back to Active to review).
  - a. Search for the batch.
  - b. Select **Update** to review the batch.
  - c. Select (for example, **SemanticModel (Capacitors)**) from the **Display Format** list.

---

**Note:** If this does not exist, select **More** and search for **SemanticModel**. Select the correct **SemanticModel** (in the case where there are multiple and press **Select**).

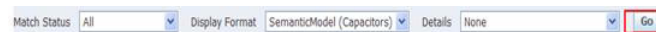
---



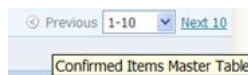
It then appears in the list as follows:



6. Select **Go** to apply it.



7. Select **Previous** or **Next** to review the items for correctness



The Standardized items are in place within the Oracle Interface Table for the selected rows from the Governance Studio job.

## 4. Reviewing the Alternate Catalog

The Governance Studio **Alternate Catalog Review** tab provides an opportunity to review and process the Alternate Catalog assignment at the item (row) level.

1. Select the **Alternate Catalog Review** tab in the Governance Studio.
2. Select from one to all of the rows in the tab.

	interface_table_unique_id	organization_id	batch_id	item_number	Alt_Catalog_Name	Alt_Catalog_Category
<input checked="" type="checkbox"/>	67643	204	10003316	KDL8298	Purchasing Catalog	Capacitors
<input checked="" type="checkbox"/>	67644	204	10003316	KDL8299	Purchasing Catalog	Capacitors
<input checked="" type="checkbox"/>	67647	204	10003316	B110556	Purchasing Catalog	Capacitors
<input checked="" type="checkbox"/>	67656	204	10003316	B110556	Purchasing Catalog	Capacitors
<input checked="" type="checkbox"/>	67663	204	10003316	KDL8301	Purchasing Catalog	Capacitors
<input checked="" type="checkbox"/>	67665	204	10003316	KDL8303	Purchasing Catalog	Capacitors
<input checked="" type="checkbox"/>	67670	204	10003316	KDL8308	Purchasing Catalog	Capacitors

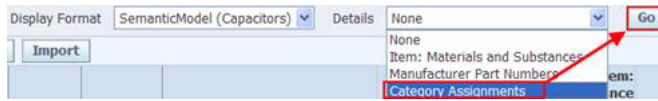
3. Click **Apply Checked Rows** button to process the selected rows.

The selected rows will turn blue, letting you know that the rows have been processed.

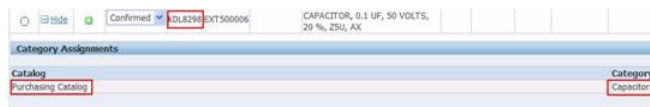
The rows become unavailable for further processing.

4. Once the processing is complete, run Release Batch process to reset the batch status from "Pending" to "Active" to review and confirm the changes made during this process. The job uses the batch id as input.
5. Confirm that the Catalog and Category are set correctly for the selected row(s) by going back into batch in the Oracle Item Import Workbench.
  - a. Search for the batch.
  - b. Select **Update** to review the batch.

- c. Select (for example, **SemanticModel (Capacitors)**) from the **Display Format** list and select **Go** to apply it.
- d. Select **Category Assignments** from the **Details** list.
- e. Select **Go** to apply it.



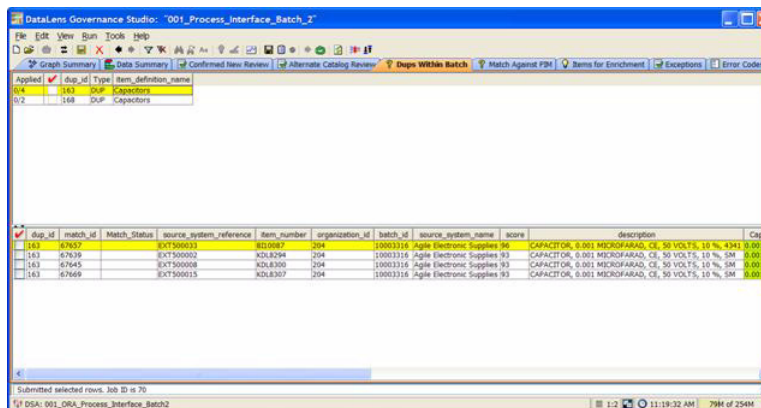
- f. Click **+Show** to open the row.
- g. Confirm that the Catalog name (for example, "Purchasing Catalog") and Category (for example, "Capacitors") is in place for all of the rows you selected earlier in this section.



The Alternate Catalog is in place within the Oracle Interface Table for the selected rows from the Interface Batch from the Governance Studio job.

## 5. Reviewing Duplicates Within the Batch

The Governance Studio **Dups Within Batch** tab uses a split screen function that shows a set of record groups (rows) in the top pane, which are considered duplicates within the import batch. By selecting a row from the top pane, once a row is selected from the top pane, the system highlights the corresponding matched candidates for that row on the bottom pane.



The duplicate rows are showed in the bottom pane. The system also gives Data Stewards color coding to highlight key attributes participating in the match. The system will color the attributes based on the type of match rule that has been triggered on the attribute. Following is the breakdown of the color coding:

- **Green Attributes:** Denotes attributes that match to the parent row and that these attributes were set to required in the item specific match rules.
- **Blue Attributes:** Denotes attributes that match to the parent row and that were set to participating but not required in the item specific match rules.



- **Blank and un-highlighted:** Indicates that the matched row did not contain a value for that particular attribute and therefore that attribute did not participate in the match score.
- **Populated and un-highlighted:** Indicates that the matched row contains a value for that attribute but that the attribute value for the matched row did not match the attribute value for the parent row.

The bottom child pane view allows you to choose either a single override row or create a combined value (blended) row by choosing a primary row with values from other rows. This survivorship process ends up with the one row or combined row that represents the set of values to be input into the system as the master record. The following steps should be used to create the master record:

1. Select the override row by right-clicking on the row and selecting **Create Override Row**. After the selection, the following will occur:
  - a. The override row will turn orange and the new master row will appear at the bottom on the set in red with a Match Status of "MO" (Manual Override). The override row check box will be selected and disabled.
  - b. All other rows will be automatically selected, but available to uncheck to exclude from processing.

dup_id	match_id	Match_Status	source_system_reference	item_number	organization_id	batch_id	source_system_name	score	description	Capac
163	47637		EXT500033	8110087	204	10003316	Agile Electronic Supplies 96		CAPACITOR, 0.001 MICROFARAD, CE, 50 VOLTS, 10 %, 4341	0.001 u
163	67639		EXT500002	KDL8294	204	10003316	Agile Electronic Supplies 93		CAPACITOR, 0.001 MICROFARAD, CE, 50 VOLTS, 10 %, SM	0.001 u
163	67643		EXT500008	KDL8300	204	10003316	Agile Electronic Supplies 93		CAPACITOR, 0.001 MICROFARAD, CE, 50 VOLTS, 10 %, SM	0.001 u
163	67669		EXT500015	KDL8307	204	10003316	Agile Electronic Supplies 93		CAPACITOR, 0.001 MICROFARAD, CE, 50 VOLTS, 10 %, SM	0.001 u
163	47637	MO	EXT500033	8110087	204	10003316	Agile Electronic Supplies 96		CAPACITOR, 0.001 MICROFARAD, CE, 50 VOLTS, 10 %, 4341	0.001 u

2. Optionally, right-click an individual cell from any of the other rows to create an override cell by right-clicking from the desired cell and selecting "Set Override Cell". The override cell will turn orange and that row will be automatically selected (if unselected) and disabled while the cell in the original override row will be deselected and return to the original color. Finally, the value will appear in the master row.

**Note:** if you want to enter a value, you must first enter the value prior to selecting that cell as an override cell; otherwise the change will *not* be reflected in the master row.

Capacitance	Dielectric	Voltage	Tolerance1	Capacitor_package_size
0.001 uF	CE	50 VLT	10 %	4341
0.001 uF	CE	50 VLT	20 %	
0.001 uF	CE	50 VLT	10 %	
0.001 uF	CE	50 VLT	10 %	
0.001 uF	CE	50 VLT	20 %	4341

You can also change the override cell selection by right-clicking an orange highlighted cell and selecting another cell using the **Set Override Cell** option. You can select an unlimited number of override cells to create the blended record. To remove an override cell, you need to right-click the cell from the red row and select **Set Override Cell**. You can also remove the override row selections by right-clicking any row and selecting **Remove Override Row**. Selecting **Remove Override Row**, resets the entire duplicate set and returns the set to the original state.

Once you complete the selection process, click the **Re-Run** button to reprocess the row(s).

**Expected Result**

The Override row (orange / red row) is reprocessed as a new row and could end up back in any of the tabs, but it is very unlikely that it will end up in the **Enrich** or **Exception** tabs because in order to participate in the Dedup process, the QI needed to be above the threshold. The other selected rows in the set, which Match Status is not changed by you, should be updated in the interface tables with a match status of "Ignore".

You may also override the system by entering a match status and sending the rows to be reprocessed. They may be one of the following:

- **Confirmed New:** After reviewing the information, you may conclude that although the system thinks these are duplicates, they are not. The row(s) will be included as new items in the PIM.
- **Unconfirmed:** (default value) - You must confirm the match status in order to import the item.
- **Ignore / Blank:** Temporarily hold this item from import. The next time the same item data is uploaded within a batch, it appears in the **Excluded** tab for that batch.
- **Exclude:** The row(s) will *not* be included as new items in the PIM and will be moved to an excluded region so that if the same item data row is loaded, it will be automatically moved to the **Excluded** tab.
- **Reprocess:** Allow you to update the data so that the row can be reprocessed and removed as a duplicate.

Once you completes the selection process, click the **Re-Run** button to process the row(s). The only instance that the data in the PIM will actually be enriched is when you select the "Confirmed New" Match Status. The system will ignore all other rows that you enriched (for example, unselected, and selected and any match type including blank).

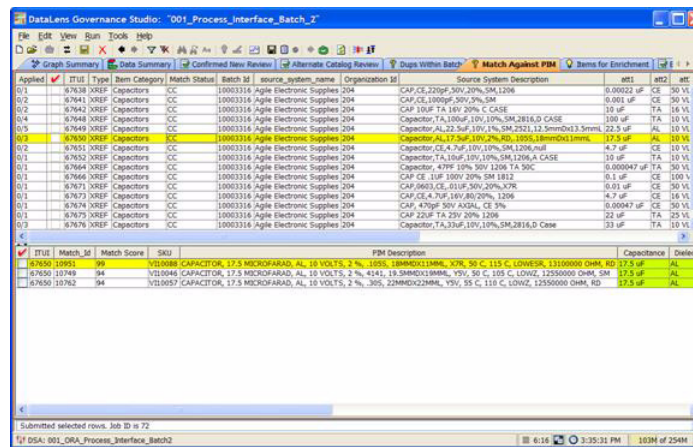
Use the following processing steps:

1. Select the **Dups Within Batch** tab in the Governance Studio.
2. Make the process selections as described previously in this section.
3. Click **Re-run** to process the selected rows.
4. For duplicate processing, confirm that the override row appears in one of the **Process Interface Batch** tabs. If the meets the quality standards and completeness, it will appear in the **Confirmed New Review** tab; otherwise, it may appear Items for **Enrichment** tab or **Match Against PIM** tab.
5. For override processing, once the processing is complete, you must run Release Batch process to reset the batch status from "Pending" to "Active" to review and confirm the changes made during this process. The job uses the batch id as input.
6. Confirm that the data is set correctly for the selected row(s) by going back into batch in the Oracle Item Import Workbench (Assumes batch set back to Active to review - see note in Introduction section).
  - a. Search for the batch.
  - b. Select **Update** to review the batch.
  - c. Select (for example, **SemanticModel (Capacitors)**) from **Display Format** list, and then click **Go**.
  - d. Select **Apply**.
  - e. Use the **Previous** and **Next** buttons to review the items for accuracy.

The override row is present on one of Process Interface Batch Governance Studio job tabs or for overridden rows in the Oracle Batch with the correct status.

## 6. Reviewing Matches Against PIM

The Governance Studio **Match Against PIM** tab uses a split screen function that shows a set of record groups (rows) in the top pane from the incoming records which are functionally equivalent (duplicate) with record(s) in the Oracle Product Data Hub. It allows the system to identify these incoming records as a cross-reference to an existing Oracle Product Data Hub record. Once a row is selected from the top pane, the system highlights the corresponding matched candidates for that row on the bottom pane. You select the row in the bottom pane that is the preferred match for the cross-reference.



The matched row(s) are showed in the bottom pane. The system also gives Data Stewards color-coding to highlight key attributes participating in the match. The system will color the attributes based on the type of match rule that has been triggered on the attribute. Below is the breakdown of the color-coding:

- **Green Attributes:** Denotes attributes that match to the parent row and that these attributes were set to required in the item specific match rules.
- **Blue Attributes:** Denotes attributes that match to the parent row and that were set to participating but not required in the item specific match rules.
- **Blank and un-highlighted:** Indicates that the matched row did not contain a value for that particular attribute and therefore that attribute did not participate in the match score.
- **Populated and un-highlighted:** Indicates that the matched row contains a value for that attribute but that the attribute value for the matched row did not match the attribute value for the parent row.

You select a row in the top pane to see a set of candidate row(s) in the bottom pane to create a cross reference row. Once the cross reference row is selected, you can send the cross reference row to the Oracle Interface Table.

---

**Note:** The ITUI (Interface Table Unique) ID / Match\_Id relationship is used for the cross reference. The Match Status is set to "Confirmed Match" which denotes a cross reference to an existing item in the PIM.

---

The system will automatically enter the PIMDH Item number (for example, VI10585) in the PIMDH Item field and will set the confirm status to "CM" (Confirmed Match).

Once you completes the selection process, you click the **Re-Run** button to process the row(s).

The cross reference appears in the batch.

---

---

**Note:** The cross reference row(s) are removed from the list and the only way to confirm is by reviewing the Oracle batch.

---

---

## 7. Reviewing Items for Enrichment

The Governance Studio **Items for Enrichment** tab with AutoSuggest output is identified by a light bulb icon to the left of the tab name. Data in these tabs may access the Governance Studio AutoSuggest feature that automatically predicts attributes that may have been missed due to misspellings and abbreviations or missing grammar. When this tab is not empty and there are suggestions available the "Suggest values for missing attributes" light bulb in the toolbar will be active; if no suggestions were found this button bar is dimmed.

After clicking on an AutoSuggest output tab, you will also see a list of rows of data that were designated for output by the respective output step in the DSA. To view suggestions, click the AutoSuggest light bulb button. Suggestions are now highlighted in green.

Several operations are allowed in the AutoSuggest table to process your data further and make it easier to edit, move, copy, or accept suggestions. In addition to the normal table functions available to output tables, an AutoSuggest table allows you to accept or reject suggestions by double-clicking over them. If you double-click a suggestion then the cell is highlighted in pink showing that the attribute has been rejected.

When you have completed reviewing the suggestions by rejecting the bad ones, keeping the good ones and editing the cells to add (enrich) the data, you may select the rows that you want to reprocess by clicking on the row then click **Re-Run**. After enriching, the item will follow the standard process and end up in one of the tabs depending on the outcome (for example, If the enrichment was sufficient then it will be of high enough quality to appear in the **Confirmed New Review** tab).

1. Select the **Items for Enrichment** tab in the Governance Studio.
2. Make the process selections as previously described in this section.
3. Click the **Re-Run** button to process the selected rows.
4. Confirm that the enriched row appears in one of the **Process Interface Batch** tabs. If the meets the quality standards and completeness, it will appear in the **Confirmed New Review** tab; otherwise, it may appear **Items for Enrichment** tab or **Match Against PIM** tab.

The enriched record is present on one of **Process Interface Batch** tabs.

## 8. Reviewing Exceptions

The Governance Studio **Exceptions** tab presents you with items that were unrecognized (did not meet the quality threshold). You has the opportunity to review and / or edit the description then make a decision about what to do with the item.

1. Select the **Exceptions** tab in the Governance Studio.

- Make a Match Status selection from the following set of selections:

IGNORE	Ignore the row
EXCLUDE	Exclude the row
CONFIRMED NEW	Confirm new row (allow into PIM)
UNCONFIRMED	Unconfirmed row that must be manually reviewed by you in the Oracle Batch
REPROCESS	Reprocess the row

- Click the **Re-Run** button to process the selected rows.
- Confirm that the row appears in the **Process Interface Batch** tab with the correct status or in one of the **Process Interface Batch** tabs if reprocessed. If the meets the quality standards and completeness, it will appear in the **Confirmed New Review** tab; otherwise, it may appear in either the **Items for Enrichment** tab or **Match Against PIM** tab.

## 9. Reviewing Style Matches

The Governance Studio **Style Matches** tab provides a view of the records that match an existing record Style Item in the PIM based on a particular set of Style attributes. The tab uses a split screen function that shows a set of record groups (rows) in the top pane from the incoming records, which match an existing Style in the Oracle Product Data Hub. Once a row is selected from the top pane, the system highlights the corresponding matched candidate for that row on the bottom pane. There should be only one Style match if the system is setup correctly. You select the row in the bottom pane that is the match.

Applied	ITUI	Batch Type	Organization Id	Batch Id	Type	Match Status	Item Category	Source System Name	Source System Reference	Source Description	alt1	alt2	alt3	alt4
Q11	266493	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1803	shirts golf cotton green checks	Green Checks Golf Cotton				
Q11	266490	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1804	shirts golf cotton orange checks	Orange Checks Golf Cotton				
Q11	266491	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1805	shirts golf cotton pink plain	Pink Plain Golf Cotton				
Q11	266492	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1806	shirts golf cotton black stripes	Black Stripes Golf Cotton				
Q11	266493	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1807	shirts golf cotton white checks	White Checks Golf Cotton				
Q11	266494	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1808	shirts golf cotton white stripes	White Stripes Golf Cotton				
Q11	266495	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1809	shirts golf cotton yellow plain	Yellow Plain Golf Cotton				
Q11	266497	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1811	shirts golf cotton yellow stripes	Yellow Stripes Golf Cotton				
Q11	266499	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1813	shirts golf cotton blue stripes	Blue Stripes Golf Cotton				
Q11	266501	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1815	shirts golf cotton red checks	Red Checks Golf Cotton				
Q11	266502	EGO_ITEM	204	10004095	STYLE_CONFIRMMATCH	Shirts	Agile Electronic Supplies	SHRT1816	shirts golf cotton red stripes	Red Stripes Golf Cotton				

ITUI	Match_Id	Match Score	SKU	Standardized Description	Color	Patterns	Shirt_use	Shirt_material
266493	234662	80		Golf Shirt, Golf, Cotton			Golf	Cotton

The matched row is shown in the bottom pane. The system also gives Data Stewards color coding to highlight key attributes participating in the match. The system will color the attributes based on the type of match rule that has been triggered on the attribute. Below is the breakdown of the color coding used on this tab:

**Green Attributes:** Denotes attributes that match to the parent row and that these attributes were set to required in the item specific match rules.

You select a row in the top pane to see a set of candidate row in the bottom pane to apply the Style. Once the Style row is selected, you can send the Style row to the Oracle Interface Table.

**Note:** The ITUI (Interface Table Unique) ID / Match\_Id relationship is used for the match cross reference. The Match Status is set to "Confirmed Match" which sets the Style information for that record.

Once you completes the selection process, you click the **ReRun** button to process the selected row(s).

The item will get an internal field called `style_item_id` set to the `match_id` in the batch.

---

**Note:** The cross reference row(s) are removed from the list and the only way to confirm is by reviewing the item in PIM.

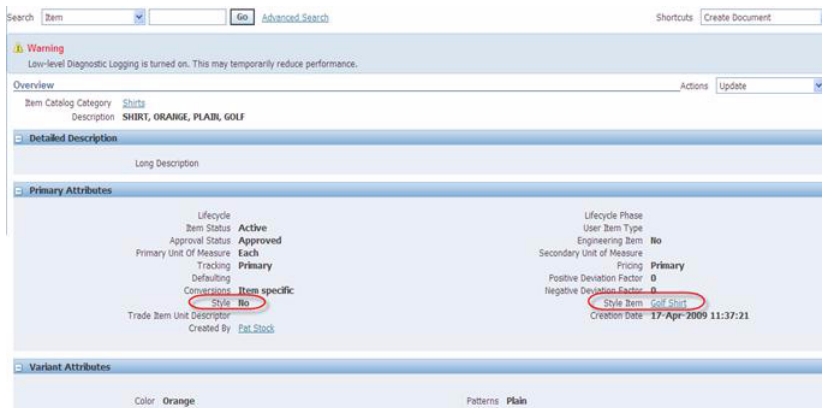
---

After the matches are applied, the processed rows will be disabled and turn blue. Once the row(s) are imported, the changes can be verified by reviewing the items have been linked to the matched style as shown below.

---

**Note:** There will be no visible indication in the Import Workbench UI about the Style association.

---



## 10. Reviewing the Release Batch

The Governance Studio **Release Batch** tab allows you to reset the batch from "Pending" to "Active" so the Oracle batch can be reviewed and imported. It also inactivates the current Governance Studio job to save the contents and prevent changes.

1. From the Governance Studio, select the **Release Batch** tab.
2. Click on the row.
3. Click on the **ReRun** button to process the selected row.

It will complete the current batch in the Governance Studio and set the project to read only (disabling most of the interface functions) and set from "Pending" to "Active".

To reactive the batch, from the File menu, select **New Project From** to reopen and process using a version number and the batch number as input.

## 11. Reviewing Error Codes

The **Error Codes** tab surfaces any errors from running the job. A typical error would be trying to run the job against a batch with an Active status when it should have been a "Completed" status.

---

---

## Installing the Client Software

Oracle Product Data Quality uses a concept called Java Web Start to initially install and maintain the current version of the software on your client desktop. The process requires you to access the Oracle DataLens Server to initiate the connection and download the software.

The Oracle Product Data Quality client applications downloaded and installed using Java Web Start by browsing to the installation page for your Oracle DataLens Server as follows:

1. Ensure that you have the Java SE Runtime Environment (JRE) 6 Update 21 installed. You can download the JRE and obtain the installation instructions by browsing to:

<http://www.oracle.com/technetwork/java/javase/downloads/index.html>

2. Start Microsoft Internet Explorer.
3. Initiate a connection and download the client software by browsing to:

`http://server:2229/datalens/datalens.`

Where *server* is the hostname of the Oracle DataLens Server.

---

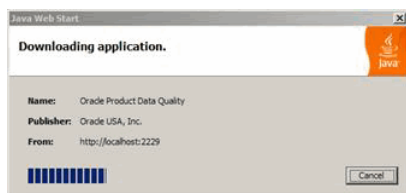
---

**Note:** If you have setup a different port number for your application server other than 2229, you must use that port number in the following URL when browsing to the Oracle DataLens Server to download the client applications.

---

---

The application download and verification begins.



---

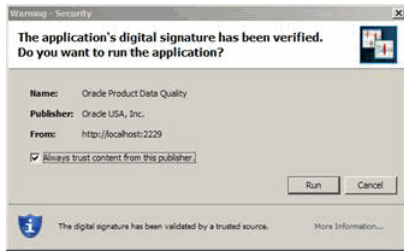
---

**Note:** If you receive a **File Download** message indicating that the . file is not associated with a program, you do not have the supported JRE installed. Click **Cancel** and return to Step

---

---

After the verification completes, the installation begins. Oracle Product Data Quality files are digitally signed by a trusted source so the following security warning is displayed:



**Tip:** To avoid this security dialog in the future, select the **Always trust content from this publisher** check box.

4. Click **Run** to continue and complete the installation.

The Oracle Product Data Quality log on dialog is displayed.

