Data Governance 3

Using Apache Atlas

Date of Publish: 2018-04-01



Contents

Searching and Viewing Entities	
Using Basic and Advanced Search	
Using Basic Search	
Using Advanced Search	
Save searches	
View entity data lineage and impact	14
View entity details	16
Manually create entities	19
Working with Atlas Classifications	21
Create Atlas classifications.	
Associate classifications with entities	
Propagate classifications to derived entities	
Search for entities associated with classifications	
Managing Business Terms with Atlas Glossaries	32
Overview	
Create a glossary	
Create glossary terms	
Associate glossary terms with entities	
Associate glossary terms with related terms	
Search for entities associated with terms	
Create glossary categories	40
Assign terms to categories	42
Amadaa Adlaa DECT ADI	43
Anache Atlas REST API	4.1

Searching and Viewing Entities

Using Basic and Advanced Search

Using Basic Search

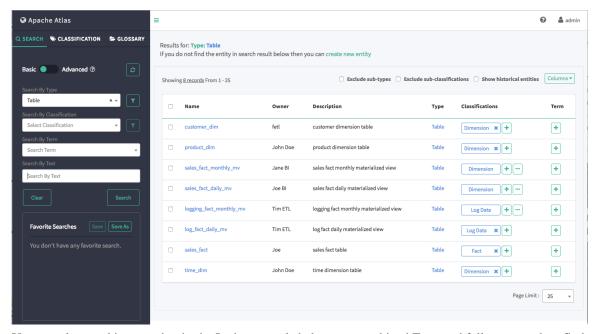
About this task

You can search for entities using three basic search modes:

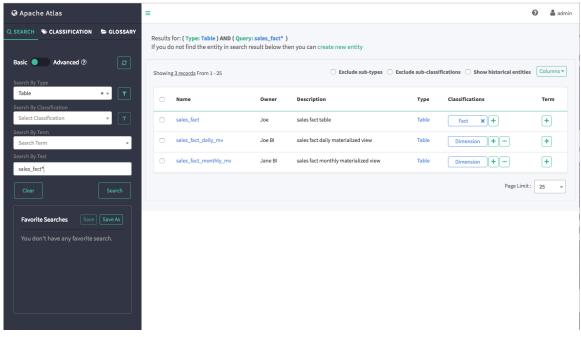
- Search by Type search based on a selected Entity type.
- Search by Classification search based on a selected Atlas classification.
- Search by Term search based on a specified glossary term.
- Search by Text full-text search.

Procedure

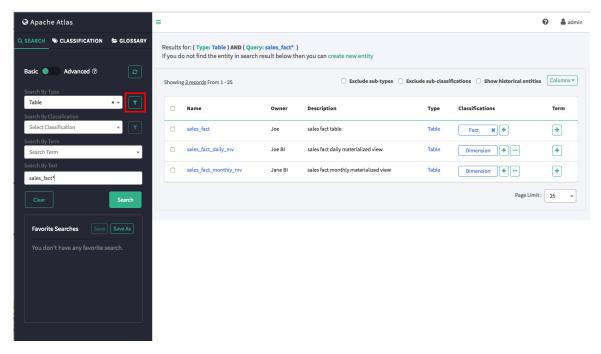
- 1. To search for entities, click SEARCH on the Atlas web UI. Select an entity type, an Atlas tag, or enter a text string, then click Search to display a list of the entities associated with the specified search criteria.
 - In the example below, we searched for the Table entity type.



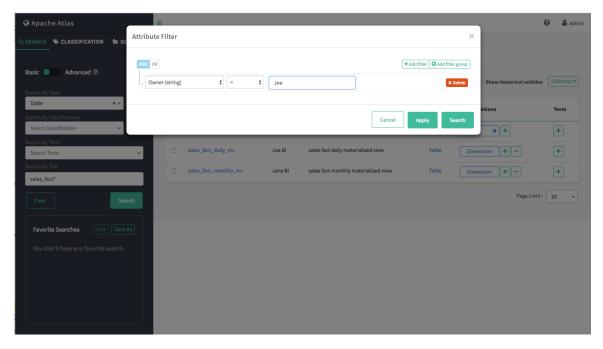
 You can also combine search criteria. In the example below, we combined Type and full-text search to find Table entities whose name contains the text string "sales_fact".



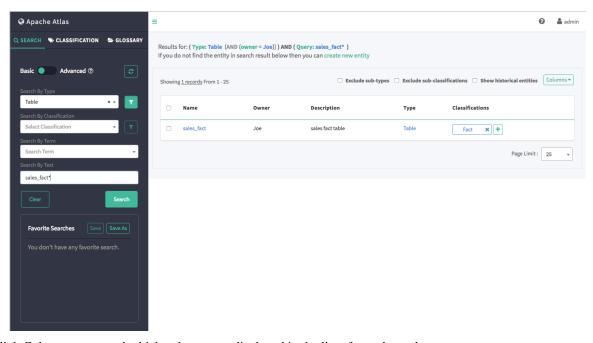
• You can use the attribute filters to further refine search criteria. Click an Attribute Filter symbol to display the Attribute Filter pop-up.



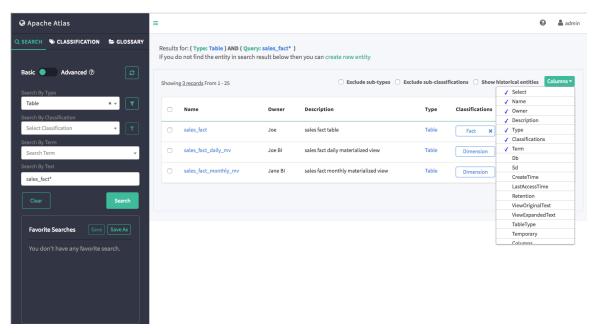
Use the selection boxes on the Attribute Filter pop-up to specify an attribute filter. The attributes listed reflect the entity type. In the following example, we set an attribute filter to return entities with an Owner attribute of "Joe".



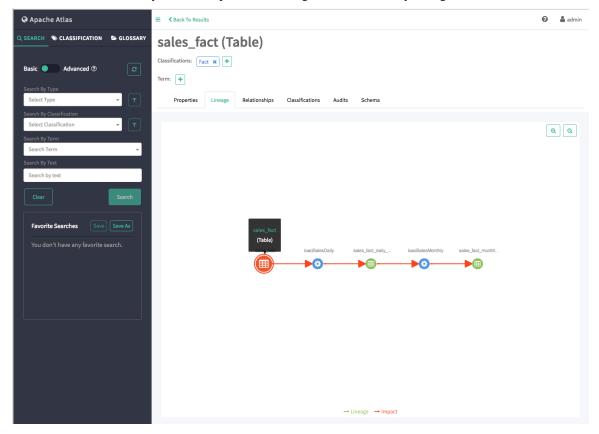
- Click Add filter to add more attribute filters.
- Click Add filter group to add an attribute filter group.
- Click Delete to remove an attribute filter or filter group.
- Click Apply to temporarily save the attribute filter to the current search without applying it to the search results. Click Search to apply the attribute filter to the search results.



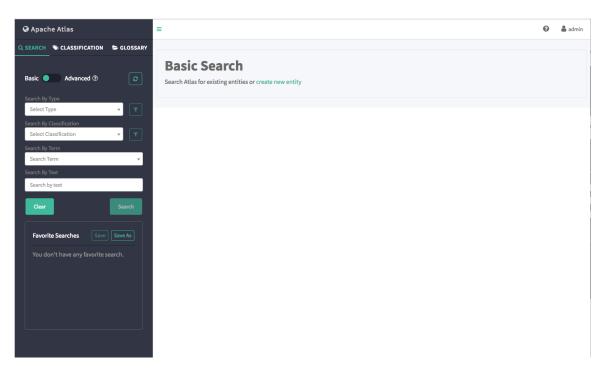
2. Click Columns to control which columns are displayed in the list of search results.



3. To view detailed information about an entity, click the entity in the search results list. In the example below, we selected the "sales_fact" table from the list of search results. You can use the tabs to view categories of information about an entity – for example, click Lineage to view the entity lineage.



4. Click Clear to clear the search settings.



Using Advanced Search

About this task

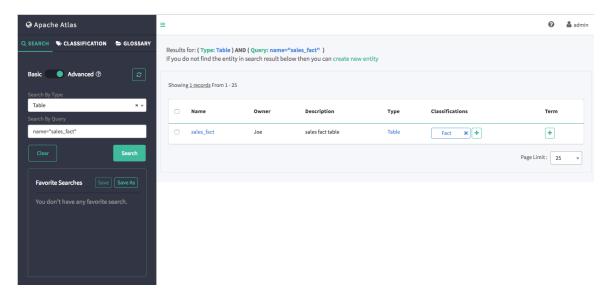
To switch to Advanced search mode, slide the green toggle button from Basic to Advanced. You can search for entities using two advanced search modes:

- Search by Type search based on a selected Entity type.
- Search by Query search using an Apache Atlas DSL query. Atlas DSL (Domain-Specific Language) is a SQL-like query language that enables you to search metadata using complex queries. For more information on Atlas DSL queries, click the Apache Atlas "Advanced Search" link under "Related information" at the end of this topic.

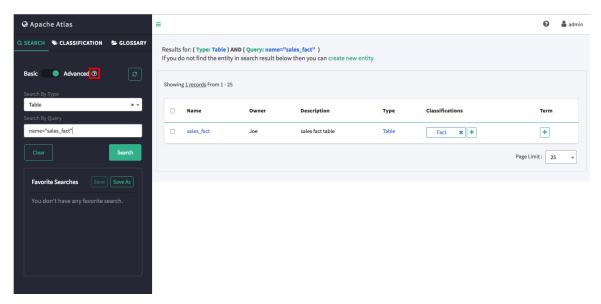
Procedure

1. To search for entities, select an entity type or enter an Atlas DSL search query, then click Search to display a list of the entities associated with the specified search criteria.

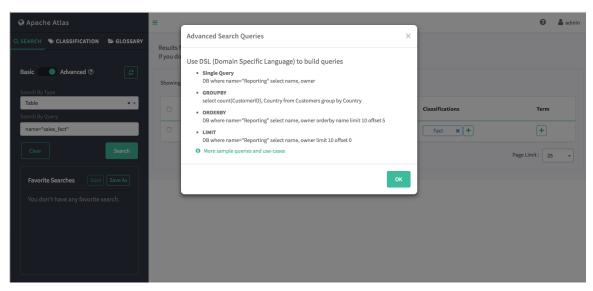
You can also combine search criteria. In the example below, we searched for Table entity types named "sales_fact".



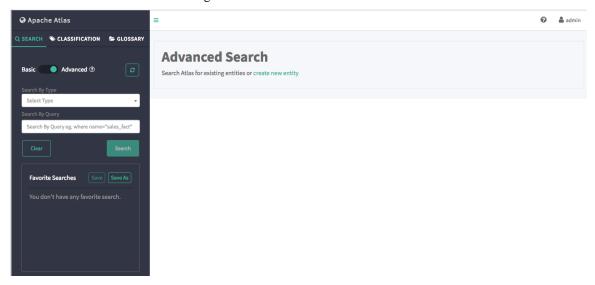
To display more information about Atlas DSL queries, click the question mark symbol next to the Advanced label above the search boxes.



The Advanced Search Queries lists example queries, along with a link to the Apache Atlas DSL query documentation:



2. Click Clear to clear the search settings.



Related Information

Apache Atlas Advanced Search

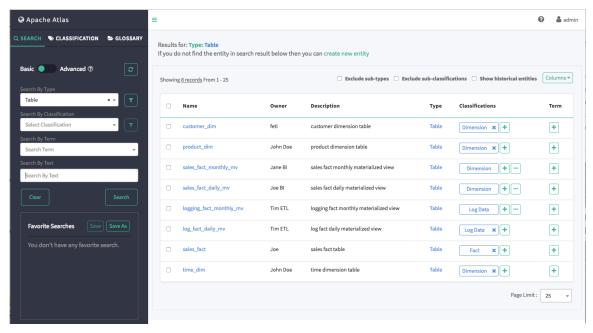
Save searches

About this task

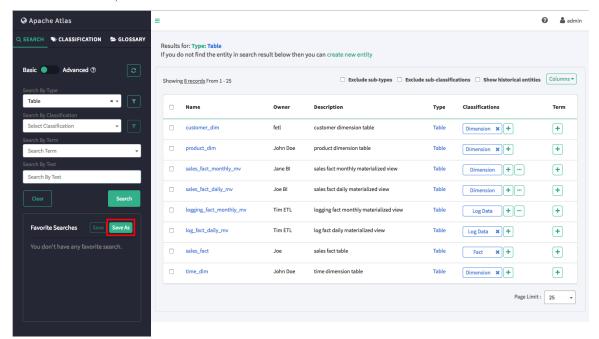
You can use the Favorite Searches box to save both Basic and Advanced Atlas searches.

Procedure

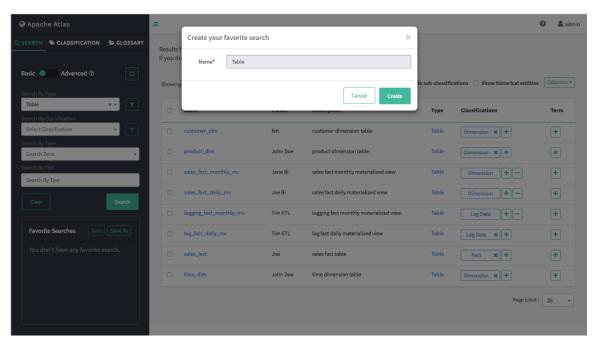
1. To demonstrate saved searches, let's start with a Basic search for the Table entity type.



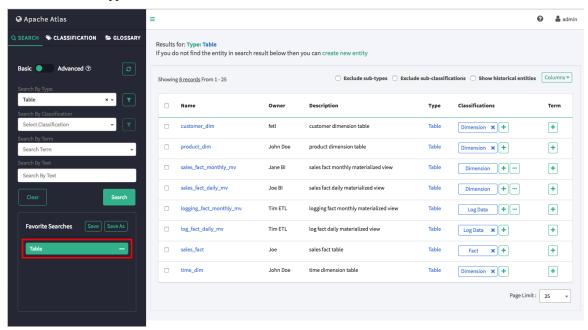
2. To save this search, click Save As under Favorite Searches.



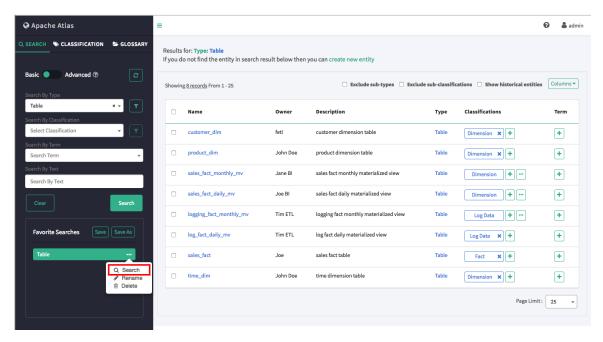
3. On the Create Your Favorite Search pop-up, type a name for the search in the Name box, then click Create. In this example, the search name is "Table".



4. The saved search appears in the Favorite Searches box.



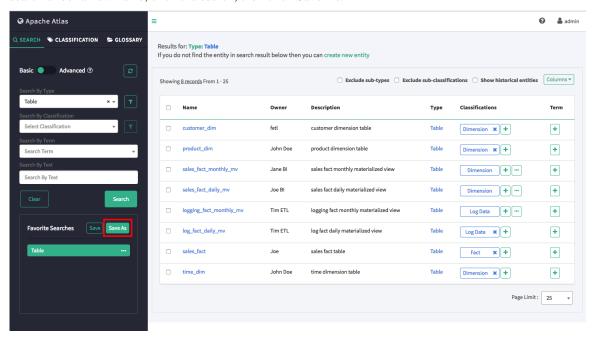
- **5.** To run a saved search:
 - Click the search name in the Favorite Searches list, then click Search.
 - Click the ellipsis symbol (...) for the saved search, then click Search in the drop-down menu.



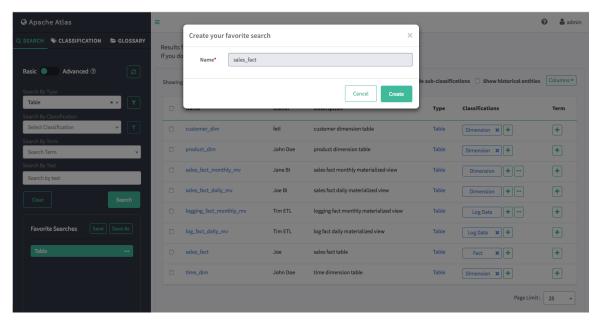
Note:

You can also use the ellipsis symbol drop-down menu to rename or delete a saved search.

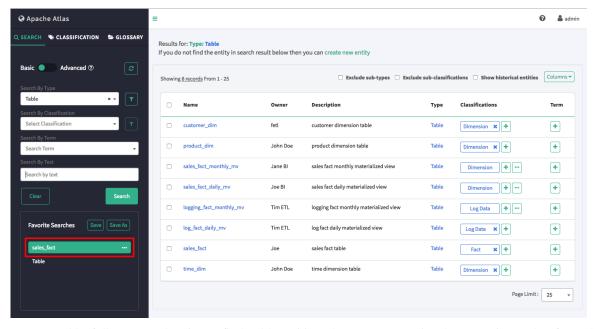
6. Now let's save this search under a new name, then change the search criteria and save the new search. To save a search under a new name, click the search, then click Save As.



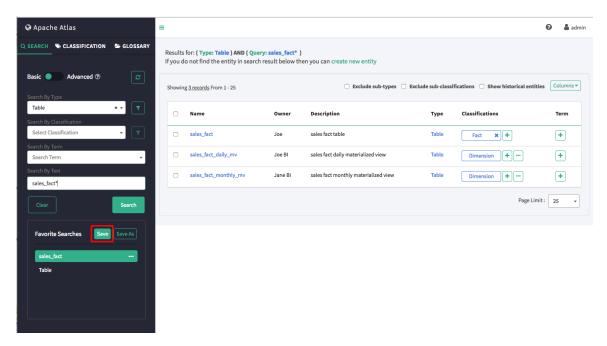
7. On the "Create your favorite search" pop-up, type a name for the new search in the Name box, then click Create. In this example, the new search name is "sales_fact".



8. The new search appears in the Favorite Searches box.



9. Next we add a full-text search string to find Table entities whose name contains the text string "sales_fact", then click Search to update the search results. To save this new set of search criteria to the "sales_fact" search, click Save under Favorite Searches.





Note:

To modify a saved search, click the search name in the Favorite Searches list, update the search criteria, then click Save to save the new search settings.

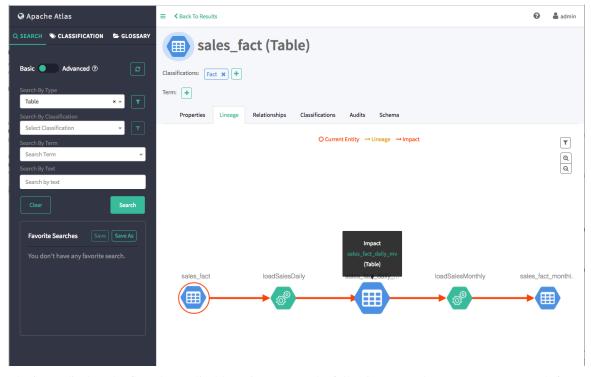
View entity data lineage and impact

Procedure

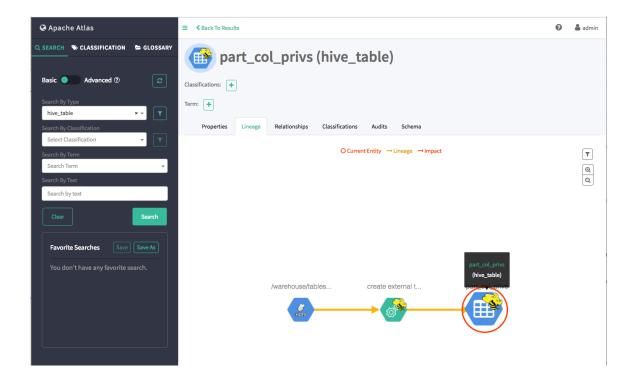
1. Data lineage and impact is displayed when you select an entity. In the following example, we ran a Type search for Table, selected the "sales_fact" entity, and then clicked the Lineage tab. Data lineage and impact is displayed graphically, with each icon representing an action. You can use the zoom buttons to zoom in and out, and you can also click and drag to move the image.



2. Moving the cursor over an icon displays a pop-up with more information about the action that was performed. Double-clicking an icon takes you to that entity's Lineage tab.



3. The icons displayed reflect the applicable entity type. In the following example, we ran a Type search for hive_table, selected the "part_col_privs" entity, and then clicked the Lineage tab. Here we can see icons for HDFS path, Hive process, and Hive table.



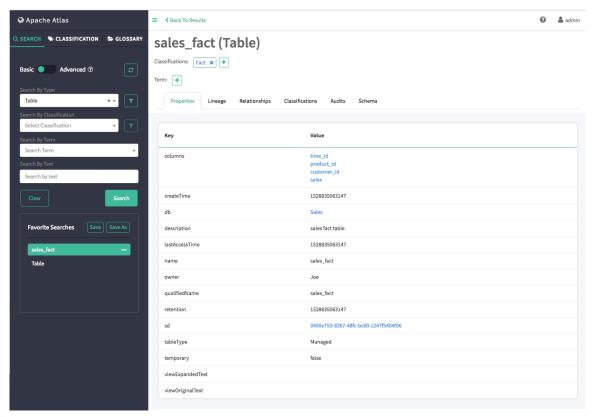
View entity details

About this task

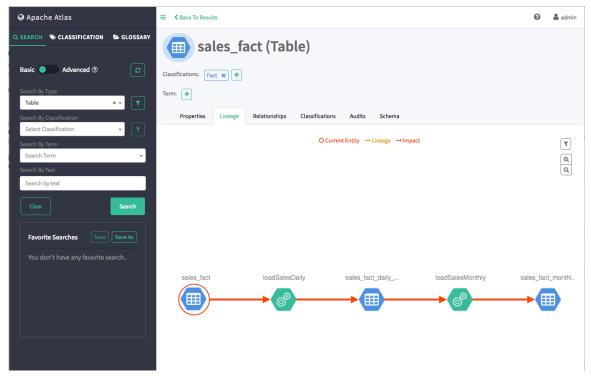
When you select an entity, detailed information about the entity is available.

Procedure

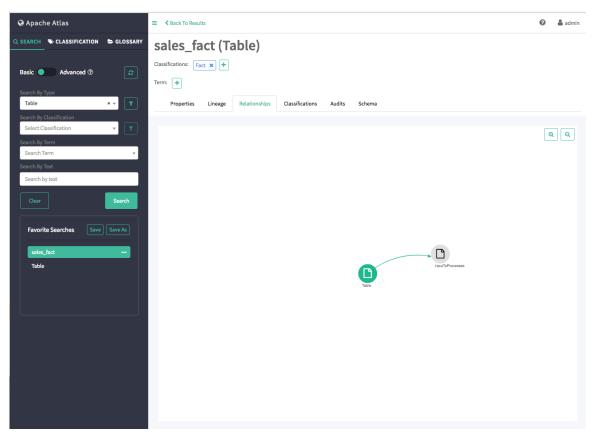
• The Properties tab displays the entity properties.



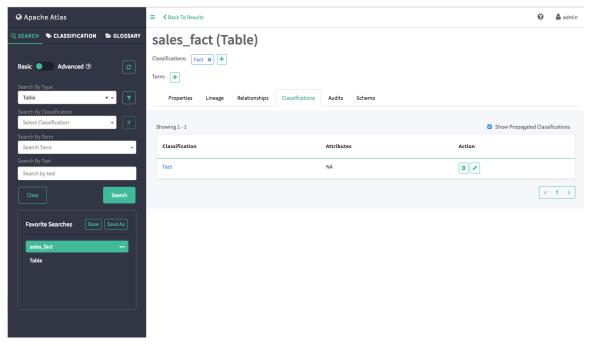
The Lineage tab displays the entity lineage.



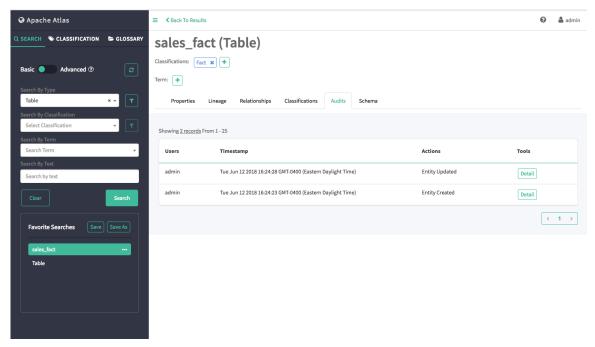
• The Relationships tab displays the entity relationships.



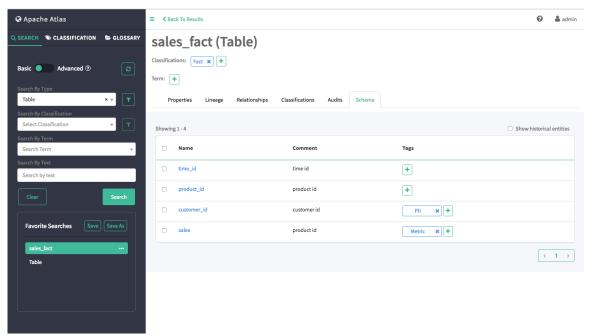
• Click the Classifications tab to display the classifications associated with the entity. In this case, the "Fact" classification has been associated with the "sales_fact" table.



• The Audits tab provides a complete audit trail of all events in the entity history. You can use the Detail button next to each action to view more details about the event.



• The Schema tab shows schema information, in this case the columns for the table. We can also see that a PII tag has been associated with the "customer_id" column.



Manually create entities

About this task

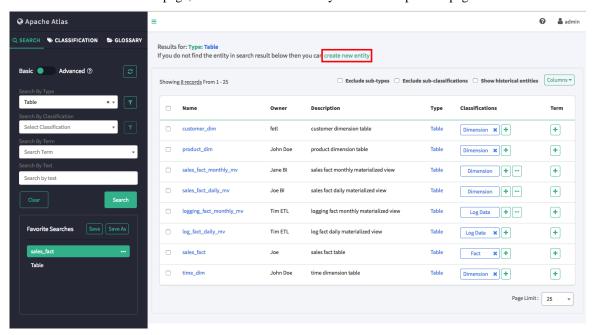
Currently there is no Atlas hook in HDP for HDFS or Kafka. For these components, you must manually create entities in Atlas. You can then associate tags with these entities and control access using Ranger tag-based policies.



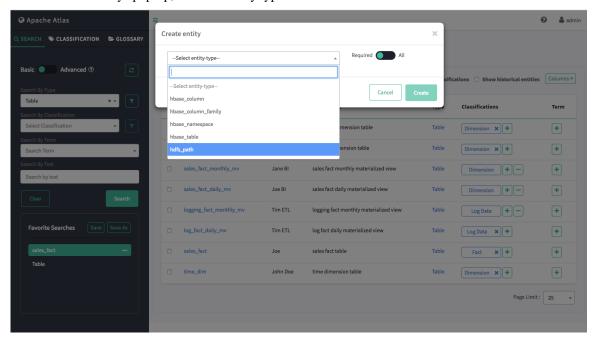
Note: When upgrading to HDP-3.0 from earlier versions, you may need to manually enable the HBase hook post-upgrade. On the Ambari dashboard, select HBase > Configs > Advanced hbase-env, then select the Enable Atlas Hook check box. Click Save, then restart HBase and any other services that require a restart.

Procedure

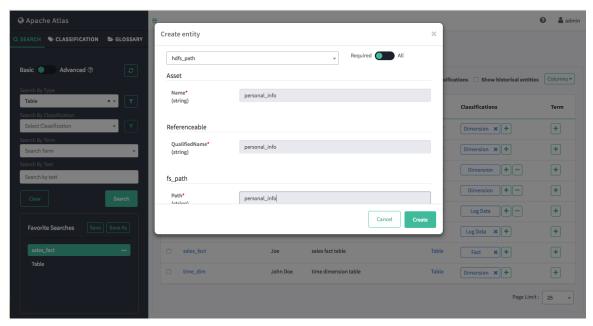
1. On the Atlas web UI Search page, click the create new entity link at the top of the page.



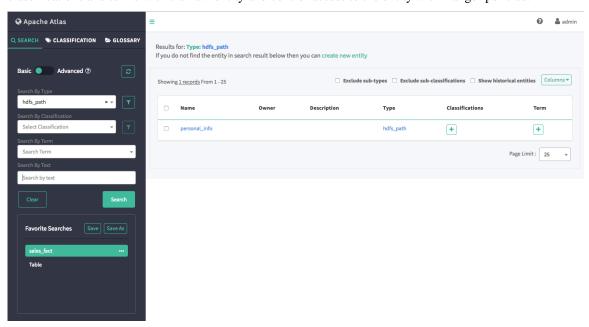
2. On the "Create entity" pop-up, select an entity type.



3. Enter the required information for the new entity. Click All to display both required and non-required information. Click Create to create the new entity.



4. The entity is created and returned in search results for the applicable entity type. You can now associate classifications and terms with the new entity and control access to the entity with Ranger policies.

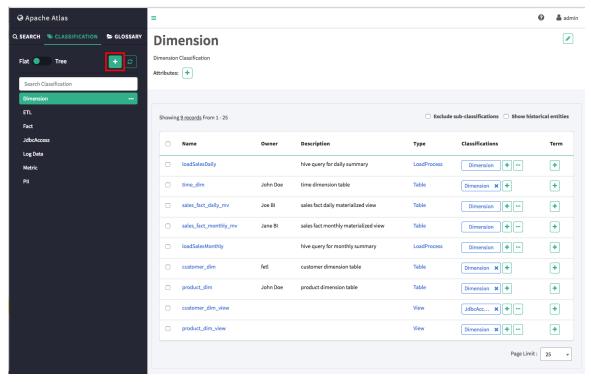


Working with Atlas Classifications

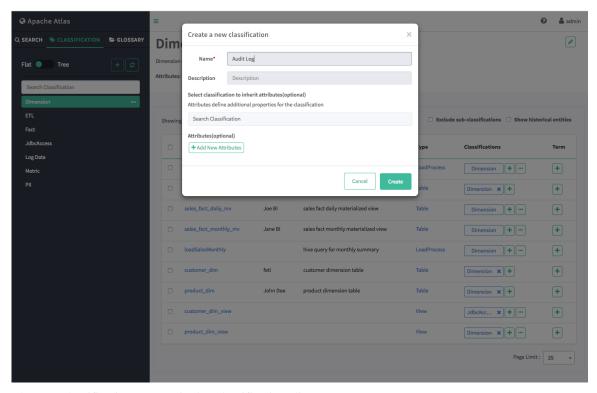
Create Atlas classifications

Procedure

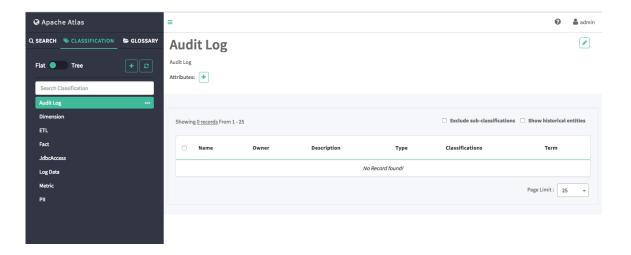
1. On the Atlas web UI, click CLASSIFICATION, then click the + icon.



2. On the "Create a new classification" pop-up, type in a name and an optional description for the classification. You can use the Select classifications to inherit attributes box to inherit attributes from other classifications. Click Add New Attributes to add one or more new attributes to the classification. Click Create to create the new classification.



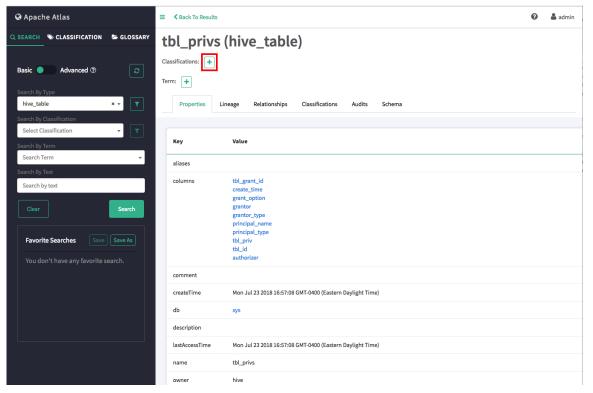
3. The new classification appears in the Classifications list.



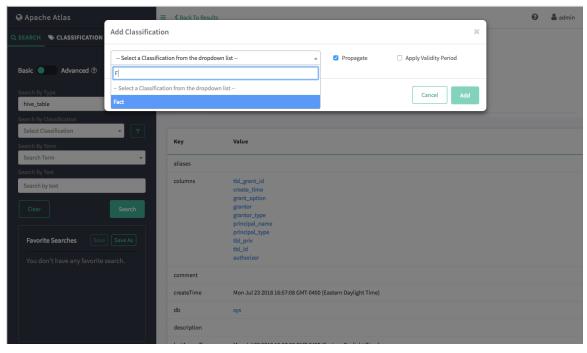
Associate classifications with entities

Procedure

1. Select an entity. In the example below, we searched for all hive_table entities, and then selected the "tbl_privs" table from the list of search results. To associate a classification with an entity, click the + icon next to the "Classifications:" label.

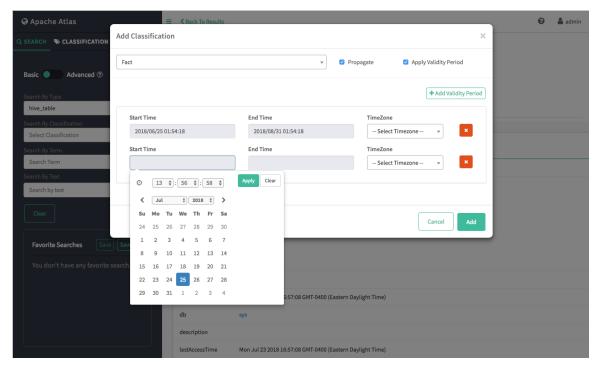


2. On the Add Classification pop-up, click the classification selection box, then select the classification you would like to associate with the entity. You can filter the list of classifications by typing text in the selection box.

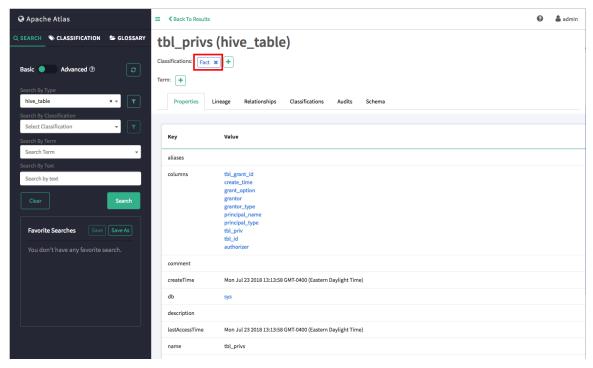


- 3. After you select a classification, the Add Classification pop-up is redisplayed with the selected classification.
 - Select the Propagate checkbox to add the classification, along with any associated attributes, to all downstream entities.
 - Select the Apply Validity Period checkbox to limit the classification assignment to a specified time period. Click Add Validity Period to specify additional validity periods.

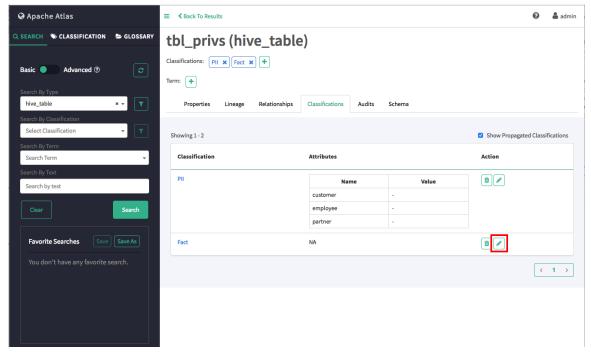
Click Add to associate the classification with the entity.



4. The new classification is displayed next to the "Classifications:" label on the entity page.

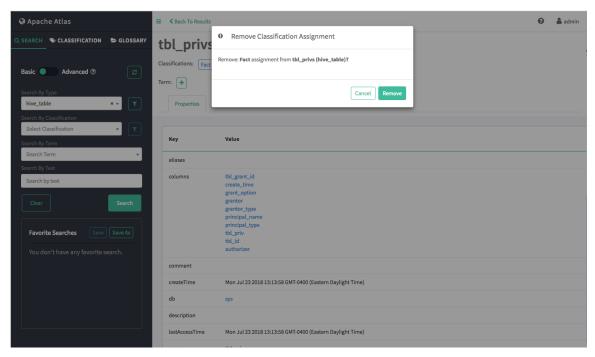


5. To edit a classification association, select the Classifications tab on the entity page, then click the Edit symbol for the classification.

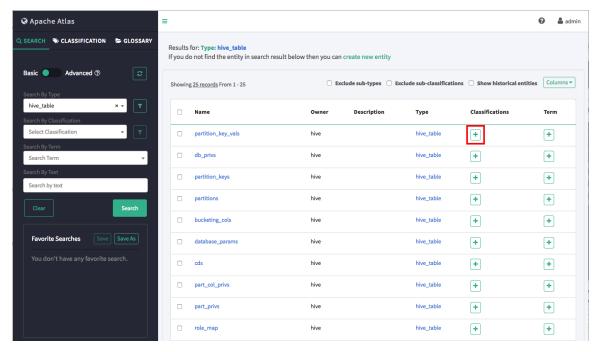


6. To remove a classification from an entity, click the x symbol on the classification label, then click Remove on the confirmation pop-up. This removes the classification association with the entity, but does not delete the classification itself.

You can also select the Classifications tab on the entity page, then click the Delete (trash can) symbol for the classification.



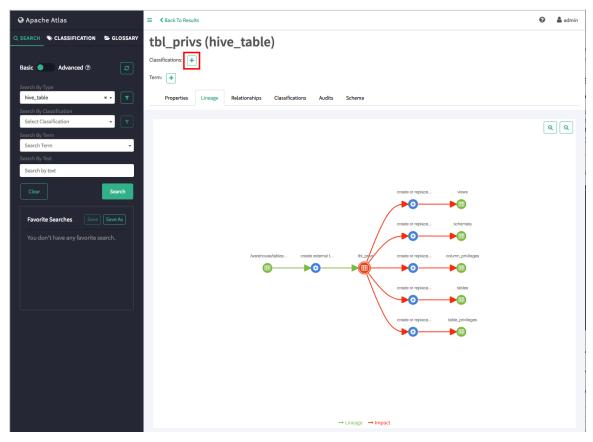
7. You can also associate a classification with an entity by clicking the + symbol for the entity in the Classifications column in a list of search results.



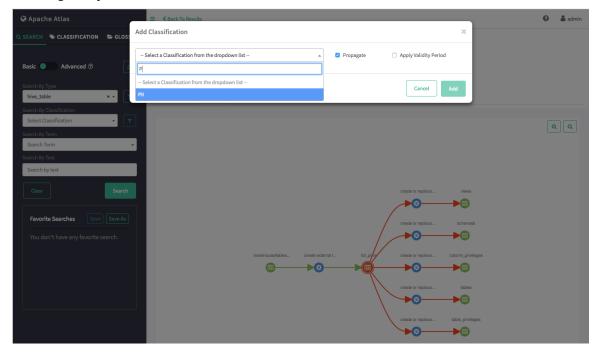
Propagate classifications to derived entities

Procedure

1. Select an entity. In the example below, we searched for all hive_table entities, and then selected the "tbl_privs" table from the list of search results. Clicking Lineage reveals that this table has derived (or "downstream") entities. To associate a classification with an entity, click the + icon next to the "Classifications:" label.

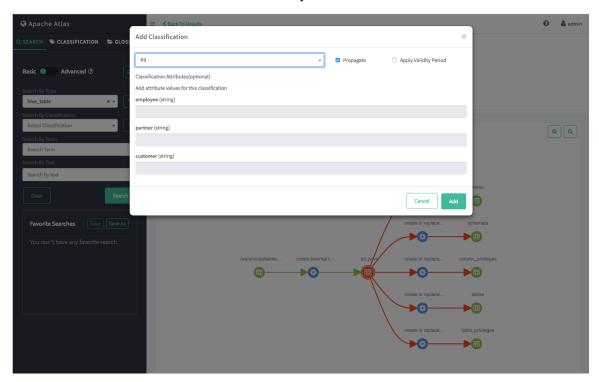


2. On the Add Classification pop-up, click the classification selection box, then select the classification you would like to associate with the entity. You can filter the list of classifications by typing text in the selection box. In the following example we select the PII classification

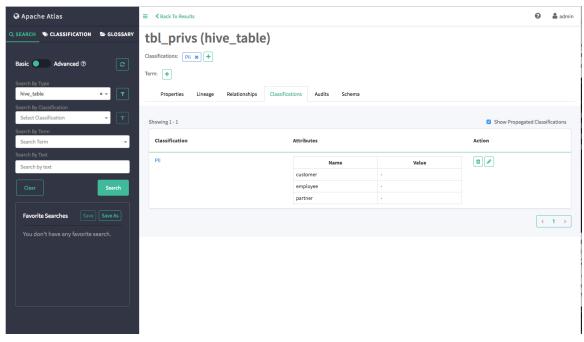


- 3. After you select a classification, the Add Classification pop-up is redisplayed with the selected classification.
 - Select the Propagate checkbox to add the classification, along with any associated attributes, to all downstream entities.

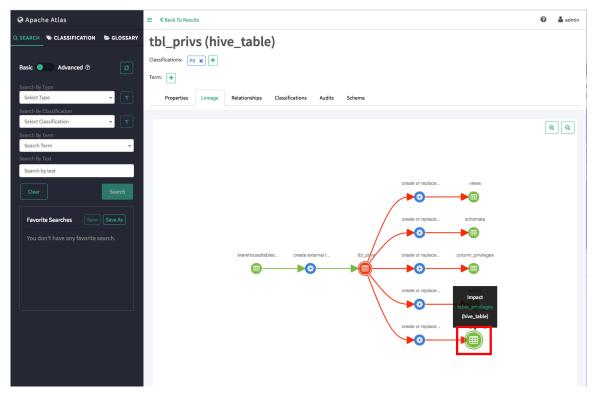
• If the classification has attributes, you can use the text boxes to add attribute values. Click Add to associate the classification with the entity.



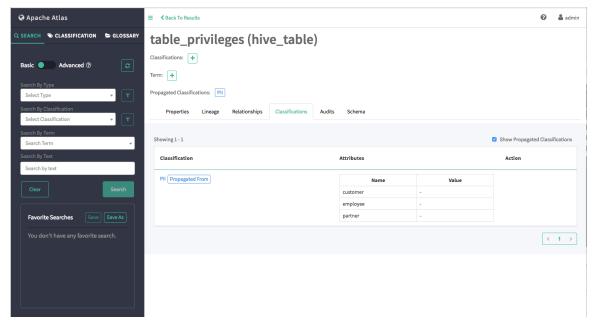
4. The new classification is displayed next to the "Classifications:" label on the entity page, and is also listed on the Classifications tab (along with any associated attributes).



5. Now we can double-click in the Lineage chart to open one of the downstream entities. In this example we double-click the icon for the "table_privileges" table.

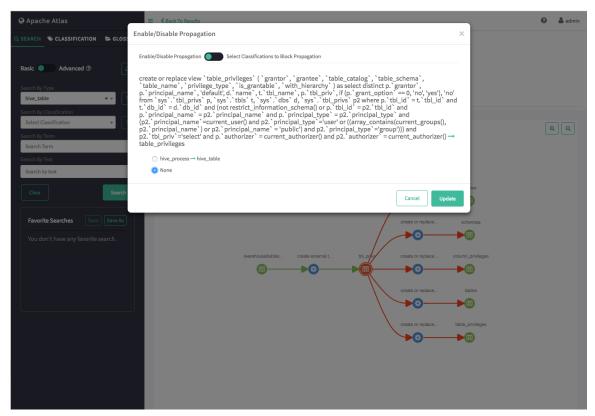


- **6.** On the "table_privileges" entity page, the propagated classification is displayed next to the "Propagated Classifications:" label, and is also listed on the Classifications tab (along with the associated attributes). Clicking Propagated From on the Classifications tab opens the "tbl_privs" parent entity page.
 - Any subsequent attribute updates in the parent entity classification will also be applied to all downstream entities.
 - Removing a propagated classification from a parent entity also removes the classification from all downstream entities.

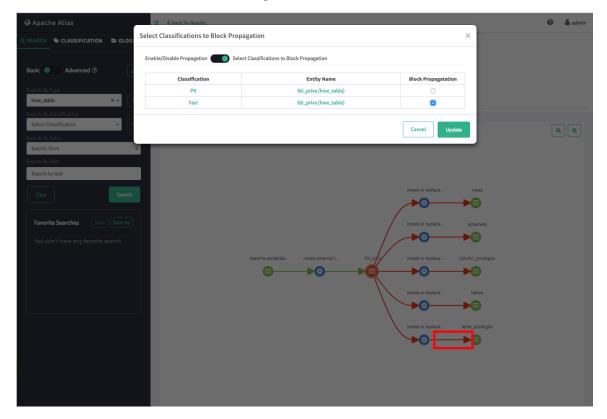


7. To remove a propagated classification from a downstream entity, click the upstream line segment ("lineage edge") next to the entity in the Lineage chart. In the following example we clicked the upstream line segment for the derived "table_privileges" tale.

• To disable all downstream classifications, select None, then click Update.



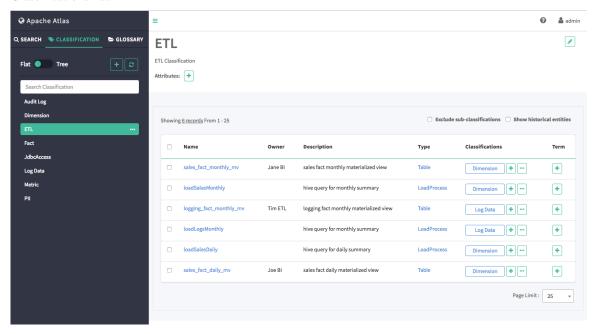
• To disable a subset of classification propagations to block, click Select Classifications to Block Propagation, select the classifications to block, then click Update.



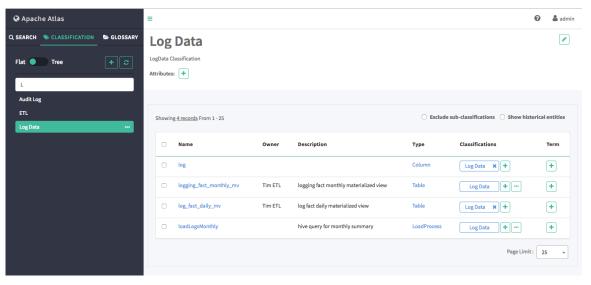
Search for entities associated with classifications

Procedure

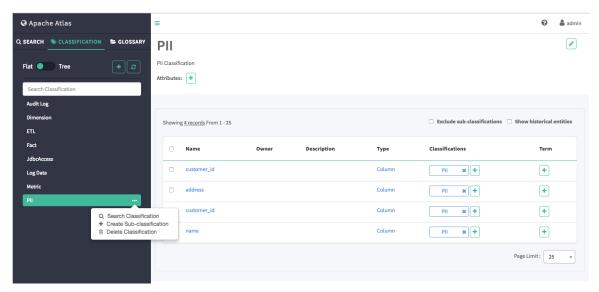
1. To display a list of all of the entities associated with a classification, click the classification name in the Atlas Classifications list.



2. To filter the Classifications list based on a text string, type the text in the "Search Classifications" box. The list is filtered dynamically as you type to display the classifications that contain that text string. You can click a classification in the filtered list to display the entities associated with that classification.



3. You can also search for entities associated with a classification by clicking the ellipsis symbol for the classification and selecting Search Classification. This launches a DSL search query that returns a list of all entities associated with the classification.



4. You can also use Basic and Advanced Search to search for entities associated with a classification.

Related Tasks

Using Basic and Advanced Search

Managing Business Terms with Atlas Glossaries

You can use the Apache Atlas Glossary feature to organize and discover data stored in Hadoop.

Overview

The Apache Atlas Glossary feature enables you to define a hierarchical set of business terms that represents your business domain. You can then associate these glossary terms with the metadata entities that Atlas manages. This hierarchical business catalog makes it easier to organize and discover data stored in Hadoop.

Atlas Terms

The Atlas Glossary Terms feature enables you to create glossaries and create terms within each glossary. The Terms list can be thought of as of a flat (but searchable) list of business terms organized by glossaries.

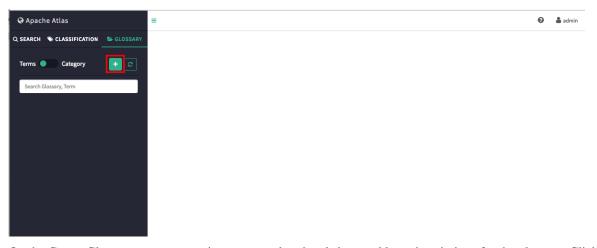
Atlas Categories

The Atlas Glossary Category feature enables you to create a hierarchy of categories and subcategories under each glossary. You can then associate terms with these categories, thereby establishing an organizational hierarchy of business terms.

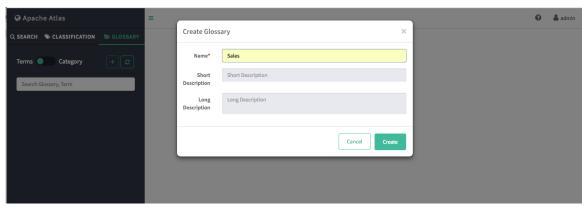
Create a glossary

Procedure

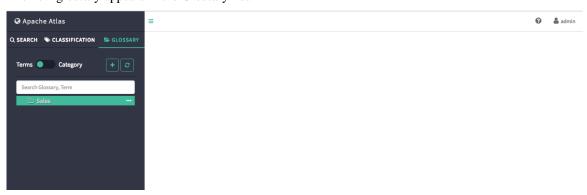
1. On the Atlas web UI, click GLOSSARY, then click the + icon.



2. On the Create Glossary pop-up, type in a name and optional short and long descriptions for the glossary. Click Create to create the new glossary.



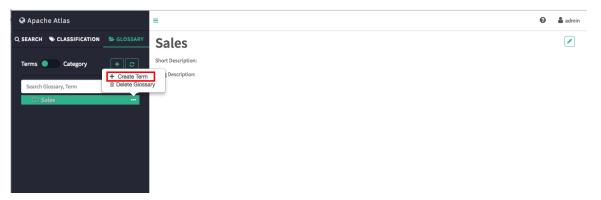
3. The new glossary appears in the Glossary list.



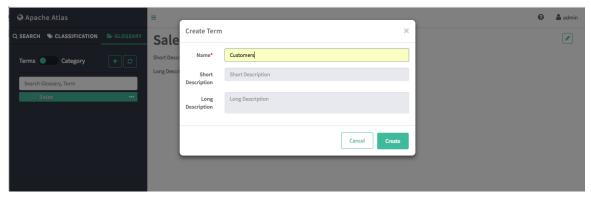
Create glossary terms

Procedure

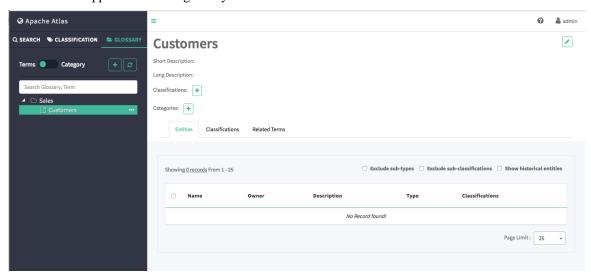
1. On the Atlas Glossary page, click the ellipsis symbol next to a glossary, then click Create Term.



2. On the Create Term pop-up, type in a name and optional short and long descriptions for the term. Click Create to create the new glossary term.



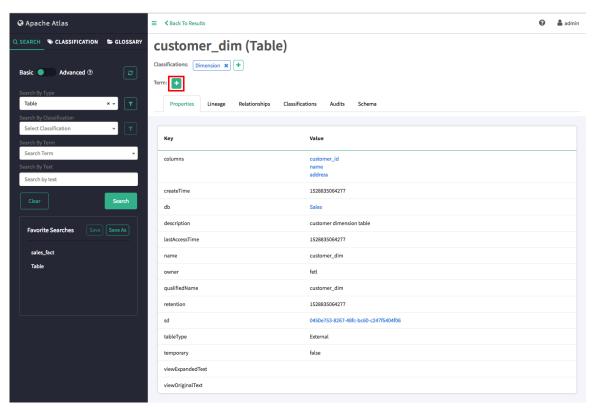
3. The new term appears under the glossary.



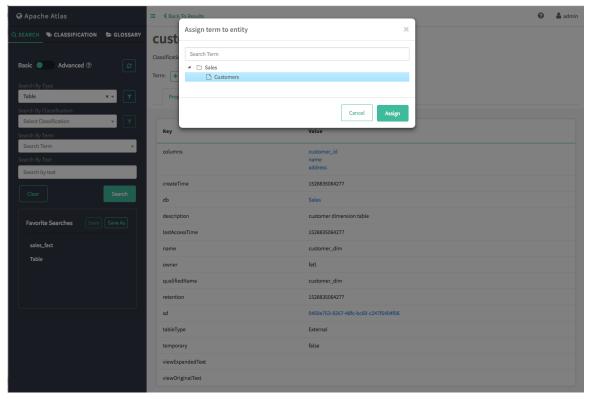
Associate glossary terms with entities

Procedure

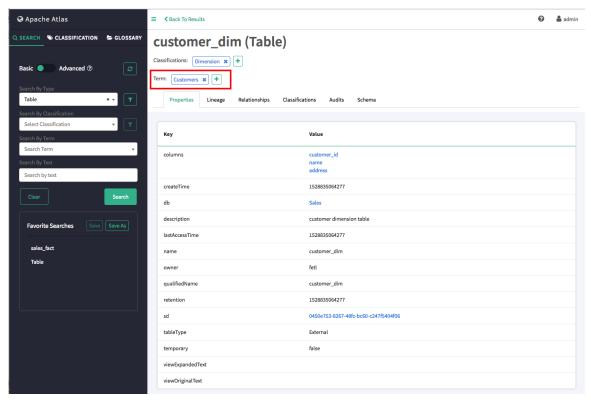
1. Select an entity. In the example below, we searched for all Table entities, and then selected the "customer_dim" table from the list of search results. To associate a term with the entity, click the + icon next to the Terms: label.



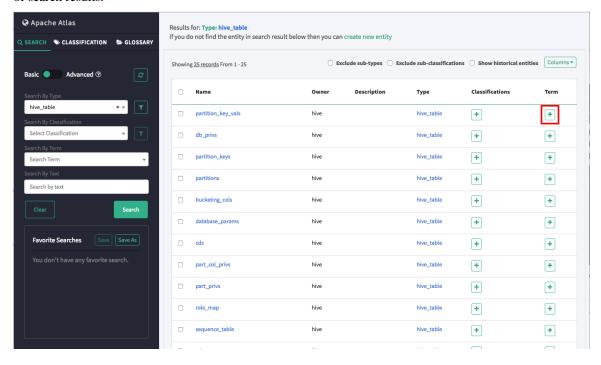
2. On the "Assign term to entity" pop-up, browse to select a glossary term. Here we have selected the term "Customers". You can also filter the list of tags by typing text in the Search Term box, and then click to select a term. Click Assign to assign the term to the entity.



3. The new term is displayed next to the Terms: label.



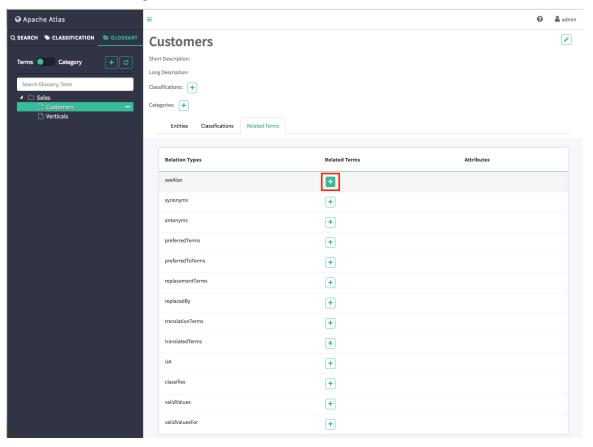
4. You can also associate a term with an entity by clicking the + symbol for the entity in the Terms column in a list of search results.



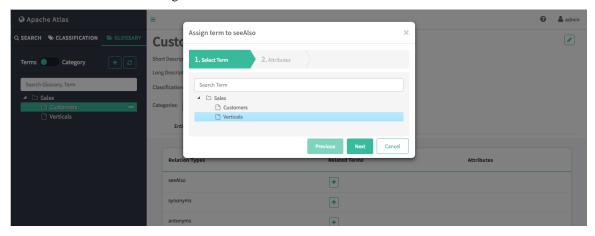
Associate glossary terms with related terms

Procedure

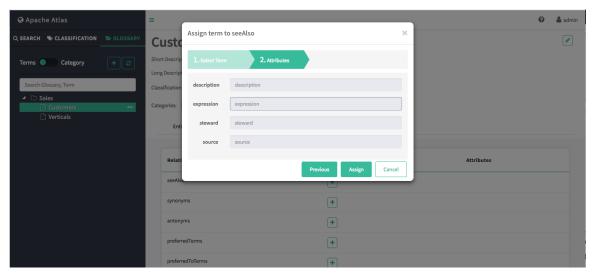
1. Select a glossary term, then click Related Terms. On the Related Terms tab, click the + symbol for the applicable related term. In this example we add a "seeAlso" related term.



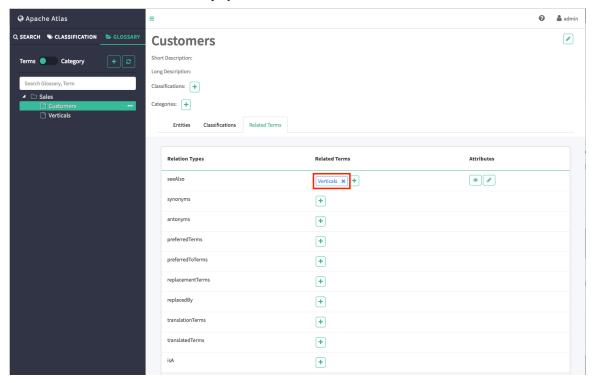
2. On the "Assign term to seeAlso" pop-up, browse to select a related glossary term. Here we have selected the term "Verticals". You can also filter the list of tags by typing text in the Search Term box, and then click to select a term. Click Next after selecting a term.



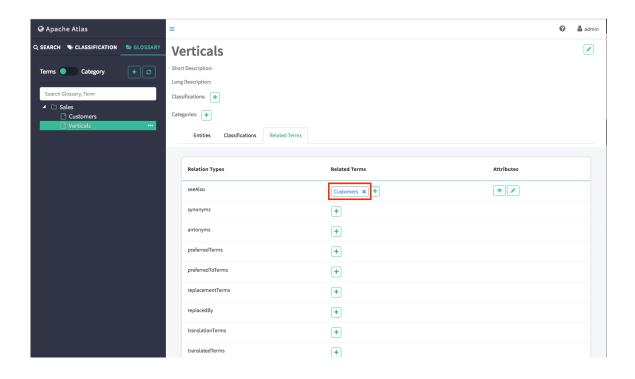
3. On the "Assign term to seeAlso" pop-up, optionally specify attributes, then click Assign.



4. The new "seeAlso" related term is displayed on the Related Terms tab.



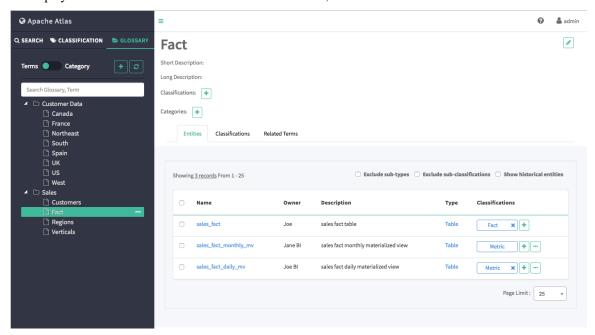
5. The "Customers" term is also listed as a "seeAlso" related term on the "Verticals" Related Terms tab.



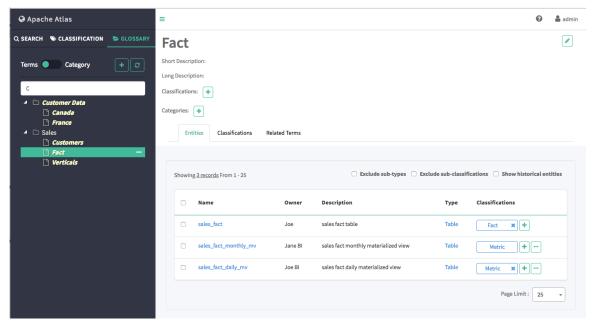
Search for entities associated with terms

Procedure

1. To display a list of all of the entities associated with a term, click the term name in the Atlas Terms list.



2. To filter the Terms list based on a text string, type the text in the "Search Glossary, Term" box. The list is filtered dynamically as you type to display the terms that contain that text string. You can click a term in the filtered list to display the entities associated with that term.



3. You can also use Basic and Advanced Search to search for entities associated with a term.

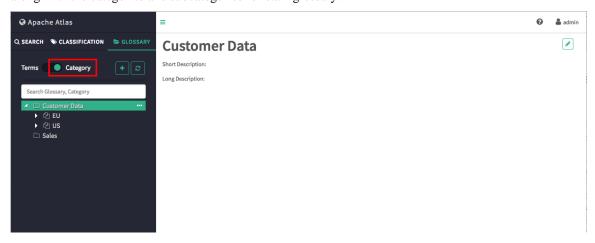
Related Tasks

Using Basic and Advanced Search

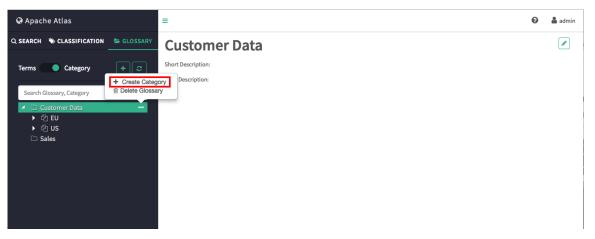
Create glossary categories

Procedure

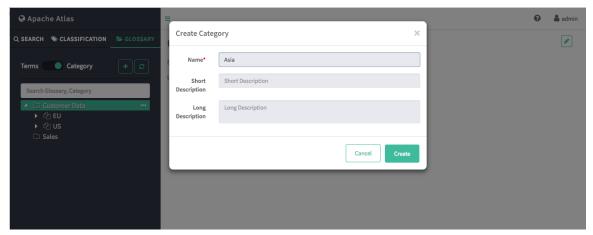
1. On the Atlas web UI, click GLOSSARY, then select Category. The Category list displays the top-level glossaries, along with the categories and subcategories for each glossary.



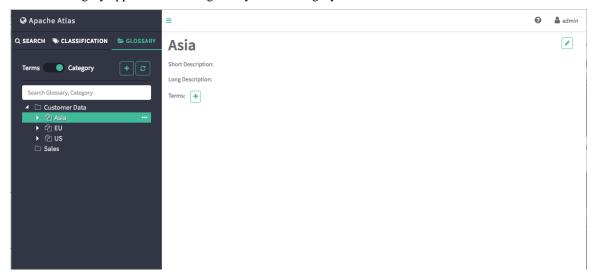
2. To create a new category, click the ellipsis symbol next to a glossary, then click Create Category.



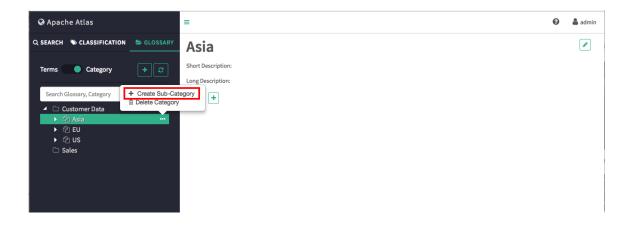
3. On the Create Category pop-up, type in a name and optional short and long descriptions for the category. Click Create to create the new category.



4. The new category appears under the glossary in the Category list.



5. To create a new subcategory, click the ellipsis symbol next to a category, then click Create Subcategory. You can create multiple levels of subcategories.



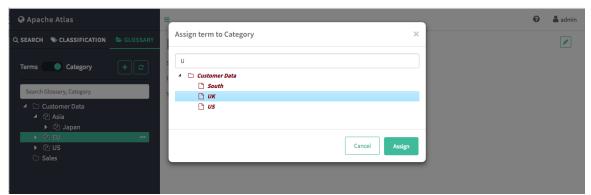
Assign terms to categories

Procedure

1. On the Atlas Glossary page, click Category. Select a category, then click the + symbol next to the "Terms:" label on the Category page.



2. On the "Assign term to Category" pop-up, select a glossary term. To filter the list of terms based on a text string, type the text in the "Search Term" box. The list is filtered dynamically as you type to display the terms that contain that text string. Click Assign to assign the term to the category.



3. The new category appears next to the "Terms:" label on the Category page.



Apache Atlas REST API

Apache Atlas exposes a variety of REST endpoints that enable you to work with types, entities, lineage, and data discovery.

Related Information

Apache Atlas REST API Apache Atlas Swagger