

# Ontologies

## Page Contents

- 1 Overview of Ontologies
- 2 Ontologies Home
  - 2.1 Prerequisites: Licensing and Enablement
  - 2.2 Listing of Ontologies
  - 2.3 Create New Ontology
    - 2.3.1 Search Indexing
    - 2.3.2 URI Construction Rules
    - 2.3.3 Includes
- 3 Using Ontologies to Configure EDG

## Overview of Ontologies

Ontologies are general data models that define classes and their properties—attributes, which store literal values and relationships that capture associations between classes.

Ontologies are special asset collections in EDG because all other asset collections are based on ontologies. In other words, ontologies define schema for data that is contained in other collections. These are either ontologies that are pre-packaged with EDG or, in case of Reference Datasets and Data Graphs, ontologies defined by users of EDG. Even for collections that are based on pre-packaged ontologies, users of EDG will commonly want to do some additional modeling - to extend and modify these ontologies.



Although an ontology can also contain *instances* of its classes, it is typically best to keep an ontology's instances in a separate model, such as in a Reference Dataset or a Data Graph. However, because ontologies *could* contain class instances, this section also includes descriptions of how to create, edit, and delete ontology instances.

TopBraid EDG supports a modular approach to defining ontologies. You can create a large single ontology model or you can create more granular ontology modules. These modules can include each other as needed. Using modules allows more granular access control. It also allows you to differentiate between entities used across the enterprise and entities used in particular parts of the business and govern them accordingly.

## Ontologies Home

Selecting the **Ontologies** link in the left-navigation pane of TopBraid EDG lists all of the Ontology collections currently available to the user and, it allows authorized users to create new ones.

## Prerequisites: Licensing and Enablement

The availability of any [collection type](#) (including Ontologies and *customer-defined* types) is determined by what is (a) licensed and (b) configured under *Server Administration*. To install a license or to view the currently licensed features, see [Setup > Product Registration](#). To configure which licensed collection types are currently enabled or disabled, see [EDG Configuration Parameters > Configure Asset Collection Types](#). For general licensing information, see the TopQuadrant website, which describes the [TopBraid products](#) and the [data governance packages](#) that determine the available collection types.

## Listing of Ontologies

This home view shows a table with all Ontologies that you can access in some way. For each collection some brief metadata is available in columns of the table. Columns are sortable and you can filter content of the table by typing search strings in the **Refine** field at the upper right of the table. To access an asset collection, click on its link.

To create an asset collection click on **Create New Ontology** button.

You can also select an asset collection in the table and start a workflow for it.

This page provides a focused view on Ontologies. To see a view of all asset collections, irrespective of their type, that you have a governance role for, click on your **User Name** in the upper right corner of the page. To see all asset collections you have access to organized by their subject areas, click on the **Governance Areas** link in the left hand-side vertical Navigation Bar.



If an Ontology is either missing or it is lacking expected features in your views, you or your security role(s) may lack proper permissions for the Ontology. A manager of the Ontology can give you the needed permissions via its utilities' **Users** settings. For background information, see [Asset Collection Permissions: Viewer, Editor, and Manager](#).

Another possible cause of a missing feature is that it requires administrative setup to become active. See [EDG Administration](#) for relevant within-application settings and/or see other [EDG Administrator Guide](#) documents for relevant external installation and integration setup.

## Create New Ontology

The **Ontologies > Create New Ontology** link opens a form with fields used to define the new Ontology. Note that you can also create a Ontology by using a Create link in the [Governance Areas](#) page.



Nobody will have a link for creating any asset collection until an administrator configures EDG's persistence technology as documented in [Server Administration: Teamwork Platform Parameters: Application data storage](#). Additionally, each user will not have a create link unless the user or their role has a *Create* permission for the EDG *Repositories* project as documented in [EDG Rights Management](#).

The **Create** dialog creates a new Ontology and automatically grants the Creator a Manager's permission for it. When Ontology creation starts from the **Governance Areas** page, new Ontology is automatically associated with the selected area. When Ontology creation starts from the Ontologies home page, new Ontology is not connected to any governance area. To change this after creation, update in utilities: **Settings > Metadata > Edit > subject area**.

The **Create** dialog box asks for the Ontology's **Label** (name), its **Default namespace** and, optionally, a **Description**. The default namespace will be used to construct URIs (unique identifiers) for the resources in the Ontology. EDG will automatically pre-populate the default namespace based on the system-wide, configurable settings. Creator can change it. Recommended practice for all collection types is to use a '/' (slash) at the end of the default namespace. For ontologies, it is typical to use '#' (pound sign). However, '/' can be used as well.

**Label**

**Default namespace**

**Description**

**Data Platform**

*Server-installed option to share this graph across all platform servers.*

## Search Indexing

If using Search the EDG with Lucene indexing (the default option), the **Create** dialog will offer an option to add this collection to the index. You can later change this setting at the **Manage** tab.

Include this asset collection in the index for *Search the EDG*

## URI Construction Rules

The **Create** dialog will also offer you an option to specify URI generation patterns for instances for each newly created graph. There are 3 different options to choose from:

New Instances Class Prefix:

New Instances Construct Method:

New Instances User Cannot Modify URI:

- **New Instances Class Prefix** has a selection of: *Default, name and acronym*. Default means that after a namespace the local name will be added. In case of name and pattern the name or pattern will be added in between the namespace and local name.

Acronym property is set for most of the asset collections in EDG, however, if you can't find acronym property in your model an extra setup needs to be done. By default ontologies don't have the `edg:acronym` property defined in a model. It requires creating a property SHAPE on an `edg:acronym` (<http://edg.topbraid.solutions/model/acronym>) property. The shape would need to associate the property with all instances of RDFS class (in a Manage tab change Root Class of Hierarchy to `rdfs:Class`, add `edg:acronym` property to `rdfs:Class` with the appropriate SHAPE defined, then change the root class to `owl:Thing` and start adding acronym to a class for which instances will be using acronym in their URI pattern)

- **New Instances Construct Method** allows to choose from: *Default, label, uuid, counter, or custom*. In this case, the local name would be either predefined by the label, uuid or counter. The Default option will be a label. Custom option allows creation of a completely custom URI (namespace and local name).

To use custom instances construct method you'd need to provide a custom method of URI creation. It requires creating a new SWP file in IDE, importing `swa.ui.ttlx`, and overwriting `'swa:createResourceDialogUsingCustom'` class accordingly.

- **New Instances User Cannot Modify URI**: *Default, true* or *false* provides the option to allow modification to the URI during the creation of assets in the asset collection. By default, Create dialog for assets will allow users to modify URI

This setup can later be changed. The option is available in the **Manage tab**. This however wouldn't change the already created instance URIs, but will take effect for the new ones.

## Includes

Collections often have natural relationships to other collections (e.g., each reference dataset's *main entity* class is defined in an included ontology). Any collection using outside resources needs to *include* the collections that contain them. Some inclusions might be required while others might merely be permitted. For example:

- Taxonomies *always* include the SKOS ontology, and they *may* include other taxonomies.
- As mentioned above, each reference dataset and data graph must include at least one ontology to define the dataset's entities.
- Glossaries always include the pre-defined EDG ontology that describes business glossary terms.
- Catalogs of data assets always include the pre-defined EDG ontology describing data assets and are expected to include definitions of relevant physical datatypes.

These requirements can be further configured. When creating a collection, any required reference to another collection will either be handled automatically or be presented for selection.

You will not be asked to provide any inclusions for a new Ontology on create.

## Using Ontologies to Configure EDG

Majority of TopBraid EDG asset collection are based on pre-specified ontologies. The exception to this rule are reference datasets, data graphs, crosswalks and content tagsets. Taxonomies are based on SKOS or SKOS-XL. All other asset collections (data assets, enterprise assets, lineage models, governance model, etc.) are based on "EDG ontologies". These ontologies are located in EDG workspace as files. Their names start with "EDG Shapes", followed by the asset collection type. For example, "EDG Shapes - Data Assets". The process for customizing these models is described below:

- Create an ontology in EDG
- Include in it an EDG Shapes model corresponding to the asset collection type you want to change
- Make changes to your ontology:
  - You can freely add classes, node shapes, properties and property shapes
  - You should not delete properties or property shapes. In fact, since they are included by reference, you will not be able to do so. Instead, de-activate a shape for a property you need to remove by setting *deactivated* to true
  - You should not (and will not be able to) remove subclass relationship between pre-built classes. Instead, you can add new subclass relationships in the ontology and then change the *Root Class* for specific asset collections (using its *Manage* tab)
- Include your customized ontology into an asset collection you want to have the modifications. If you changes should apply to all collection of a given type, EDG Administrators can define it as a default include for the type. This option is available under the EDG Configuration Parameters page.

If different groups in your organizations need to have different customizations, each should follow the process outlined above to address their specific customizations needs. [EDG Getting Started with Business Glossaries](#) guide provides an [example of using this approach](#) to customize a glossary.

If your organization decides that you need to physically remove anything in the pre-built ontologies, such changes can only be made directly to the underlying files. Should you decide to do this, the changes you make will need to be backed up and, possibly, reconciled with changes to the EDG models in the future releases.