Content Tag Sets

Page Contents

Overview of Content Tag Sets

Content Tag Sets are used for tagging content using vocabularies managed in EDG. Users can tag (assign metadata to) content through a visual user interface that displays the context for both the content and the vocabulary. They can also run Tagger's auto-classification capability to automatically assign relevant tags to content and review the results.

Content Tag Sets Home

Selecting the **Content Tag Sets** link in the left-navigation pane of TopBraid EDG lists all of the Content Tag Set 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 Content Tag Sets 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 Content Tag Sets

This home view shows a table with all Content Tag Sets 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 Content Tag Set** button.

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

This page provides a focused view on Content Tag Sets. 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 a Content Tag Set is either missing or it is lacking expected features in your views, you or your security role(s) may lack proper permissions for the Content Tag Set. A manager of the Content Tag Set 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 Content Tag Set

The **Content Tag Sets > Create New Content Tag Set** link opens a form with fields used to define the new Content Tag Set. Note that you can also create a Content Tag Set 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* .

Whenever a new Content Tag Set is created, the system will by default add an owl:import to a SHACL graph called *taggershapes*. For implementations done before version 5.2, the only way to add this would be a manual outside procedure. With this import, users can go to the Constraint Violations report and will get a screen as shown in the attachments, with an Apply button to quickly delete the triples. Basically, whenever either the subject or the object of a tag triple is deleted, an entry will be shown.

An EDG administrator sets up the content and property graph choices, as described in Configuring content and property graphs, which describes not only the setup but the information that must be in these graphs for Tagger to display them properly.

After completing this, to create a new tag set, select **Content Tag Sets** from the main EDG left-navigation menu and then select **Create New Content Tag Set**, which invokes a wizard for configuring your new tag set. The first step defines the following:

- **Label** The name that will appear for this tag set in the list of tag sets on the EDG main screen.
- **Description** A description of the tag set to help people understand what it's for.
- **Content Graph** The list of content resources to tag. A single content graph might list a set of newspaper articles or a collection of journal articles.
- **Tag Property Graph** The list of relationships choices available when tagging content resources with concept terms. For example, when assigning the concept term *Business* to a news story, a given tag property graph might offer tag property choices such as *mainSubject* and *secondarySubject*. Tag properties must be defined with several rules detailed in Tagger User Guide: Configuring content and property graphs.
- **Tagging Vocabulary** The concepts used to tag the contents. This can be either a SKOS taxonomy, or an RDFS/OWL ontology, with *skos:broader* or subclass relationships forming a concept hierarchy. A news story vocabulary would include terms such as Business, Sports, and Weather, while an academic journal vocabulary would typically have technical terms more specific to a particular academic field such as medicine or mathematics.

Content Graph, Tag Property Graph and Tagging Vocabulary are all drop-down fields that offer you a specific choice to select from. Tagging Vocabulary lists all vocabularies displayed on the system, while the selection listed on the other two drop-downs are configured by your EDG administrator as described in the section Configuring content and property graphs.

Note that there are four graphs defined in a Tag Set:

- **Content Graph:** The resources to be tagged.
- **Tag Property Graph:** A graph of properties used for tagging. The set of choices includes both administratively enabled graphs and viewable ontologies.
- **Tagging Vocabulary:** The controlled vocabulary used for tagging.
- *Content Tag Set graph:* This graph, containing all of the {<content> <property> <vocabulary-term>} triples. This is the sole data created in a Content Tag Set, as the three other graphs (content, properties and concepts) can be any other EDG project or arbitrary graph as long as they are selected as such by your EDG administrator.

After you set these values and click the Next button, the second configuration screen lets you customize the Tagger interface for this tag set with the following fields:

- **Default Tag Property** If most tagging will be done with the same property, selecting it here can make the tagging go more quickly.
- **Tag Properties** This lists all properties in the Tag Property Graph, as selected on the previous screen. Only properties with an rdfs: range value of skos:Concept will appear in this list. Only checked properties and the default tag property will appear as choices to users tagging with this Content Tag Set.
- **Root Content Type** The class/subclass structure of an RDFS or OWL model available in the content graph will appear here. One class can be chosen as the root class that will appear along with its subclasses as the **Content Types** hierarchy of EDG Tagger. Expand the tree to find the content type that you want to serve as the root and select it.

Create New Content Tag Set (Step 2)

Please select which properties can be used for tagging.

Default Tag Property

> has related ▼

Tag Properties

- ☐ has broader
- ☐ has broader match
- ☐ has broader transitive
- ☐ has close match
- ☐ has exact match
- ☐ has narrower
- ☐ has narrower match
- ☐ has narrower transitive
- ☑ has related
- ☐ has related match
- ☐ has top concept
- ☐ is in mapping relation with
- ☑ is in semantic relation with

Optionally, select a root class for the content types tree.

Root Content Type (optional)

- Resource
  - Document
    - Congress
    - Editorial
    - Journal Article
    - Letter
    - Review

☑ **Hide unused classes**

Finish

Once you have finished configuring your new Content Tag Set, it will appear as a link on the Content Tag Sets tab.