Introduction to EDG

Page Contents

TopQuadrant's TopBraid Enterprise Data Governance™ (EDG) is a flexible, web-based solution that addresses data governance needs in enterprise environments with heterogeneous data stores, data processing, and applications.

Data governance concerns how data is produced and consumed, what data is supporting business functions and enabling business capabilities, how security considerations govern access, and how policies govern needs such as data retention and data provisioning. An over-arching need is how meaning can be ascribed and traced from business glossaries to data elements. With the advent of big data, data governance increasingly needs to address how transformations execute as jobs in configurations of Hadoop stores.

TopBraid EDG supports integrated data governance across the ever growing numbers and types of data assets and governance needs. It lets you govern business glossaries, data models, conceptual models or ontologies, reference data, policy models and more. Further, it lets you build relationships across all the models - because connections are key to data governance.

TopBraid EDG is highly configurable. EDG defines a variety of vocabulary and asset types that it can govern. These definitions are model-driven. When installing EDG in your organization, the models can easily be changed or extended. Additional asset types can be defined. These guides show all of the vocabulary and asset types, but the ones available in your installation might differ: some types are license-dependent and, even when licensed, your organization may decide to initially (or permanently) configure them out.

Features of TopBraid EDG

- **An Intuitive Graphical User Interface** – with auto-completion, drag and drop, rich text editing, search and filtering, accessibility across all browsers – providing an easy-to-use environment for both business and technical stakeholders
- **Flexible data and relationship modeling** – handles both complex and simple data models and their relationships across domains; allows modeling, storing and using not only codes, but all relevant associated information
- **Auditability** – every change is logged and time stamped, change history can be searched, usage records captured where reference data is used
- **Control over versions** – virtual work-in-progress copies of reference datasets allow parallel development of versions and enable controlled publishing, review and approval workflow
- **Collaboration** – through access and accountability based on roles, support for task assignments, statuses and issues
- **Shared semantics** – providing the ability to define and share meaning of all information elements globally and in the context of specific use
- **Repeatability of on-boarding** – capturing processes and best practices for on-boarding of external reference data
- **Wide and diverse distribution** – support for a variety of interaction patterns (for example, batch or real-time) and integration approaches
- **Data quality** – offering intuitive forms for creating data validation rules
- **Integration** – with third party systems and Linked Data clouds via Web Services interfaces and APIs
- **Easy extensibility** – configurable user interfaces, reports, metamodel, import, export, web services and more including deep customizations using TopQuadrant's TopBraid platform
- **Standards** – built-in support for W3C (World Wide Web Consortium) standards for data and data model interchange on the web such as RDF and SPARQL
- **Enterprise-readiness** – scalable and robust architecture with LDAP and JMS integration

Available Asset Collection Types

TopBraid EDG supports the following vocabulary and asset model-types:

- Glossaries
- Requirements Assets
- Big Data Assets
- Data Assets
- Datatypes
- Enumerations
- Enterprise Assets
- Technical Assets
- Lineage Models
- Taxonomies
- Reference Datasets
- Data Graphs

- Ontologies
- Corpora
- Content Tag Sets
- Crosswalks

Documentation

The **Terminology** and **Getting Started** pages introduce some of the key concepts and features of EDG

The **Adminstrator Guide** : Multiple pages cover the **Installation and Integration** of the EDG application server and the **Administration** of the application's user environment and features.

The **User Guide** : Multiple pages cover the details of editing and managing the various **Asset Collections**, along with their **governance and workflows**.

The **Developer Guide** : How to customize aspects of the application, data preparation, and services, including use of the development environment *TopBraid Composer – Maestro Edition (TBC-ME)*.

Documentation Conventions

Names for user interface widgets and menu options are presented in a style like this.

Where exercises require information to be typed into EDG, a monospaced font is used.

When describing capabilities and user interfaces that work the same for ontologies and reference datasets this guide uses the term **Vocabularies** to mean both ontologies and reference datasets.

Exercises and required tutorial steps are presented like this:

1. Execute the first step.
2. Then, execute the second step.
3. Here is the third step to carry out.

Tips and suggestions for using EDG are presented like this.

Potential pitfalls and warnings are presented like this.

General notes are presented like this.