

White Paper

Maturing Information Governance with TopBraid EDG



TopQuadrant's comprehensive data governance solution, TopBraid EDG, with its semantic standards-based foundation, can manage the entire range of enterprise information assets and the cross-connections between them.

INFORMATION MANAGEMENT AND GOVERNANCE IN THE ENTERPRISE CONTEXT

Data and, in a more general sense, information is the bloodline of a modern enterprise. It is used in every decision, it is produced by every activity, and it flows across all systems. Maturity of information management is, therefore, essential to success, profitability and even survivability of every enterprise.

In recent years, organizations large and small have come to realize that maturing enterprise information management (EIM) requires establishing and maturing *information governance*. Information governance “is the oversight and management” discipline leveraging people, process, and technology to ensure information assets are suitable for use to run the business. As modern enterprises depend on technology for all day to day operations, their technology and data landscape has grown increasingly rich and complex. Most organizations can’t govern this landscape without sophisticated tools. Simply put, the complexity of today’s information landscape makes the technology part of information governance increasingly critical.

Just like the words data and information sometimes, depending on the context, are used interchangeably, the terms data governance and information governance may also be used interchangeably by some practitioners. Other practitioners may want to be clear on the difference between them. In short, information governance encompasses data governance as one of its key components.

Data governance is about creating and using policies for maximizing availability, integrity, security, and usability of

structured and unstructured information available to an organization. Information governance brings into the picture the lifecycle and business context of the information. This context includes regulatory, legal, risk, environmental, and operational requirements. TopBraid EDG supports both types of governance in an integrated way — the more tactical and detail oriented data governance together with the more strategic, business policy and context oriented information governance.

As the first semantic technology based information governance and EIM platform, TopBraid EDG provides flexible tooling to handle the complexity of information assets. It enables staff to begin in one area, shift to another, fill gaps in another, add relationships between them, and it can be done incrementally. It can bring information together through direct data entry, by importing from a set of common data formats, or through a standard API.

As a result, continued management of data and metadata within siloes, integration of the business across information siloes, and the ability to automate on top of the metadata directly — are all supported.

Figure 1, (see below), provides a view of EIM capabilities and components. The legend provides some examples of components with further descriptions, along with a brief indication of the support that TopBraid EDG provides for the respective component or capability.

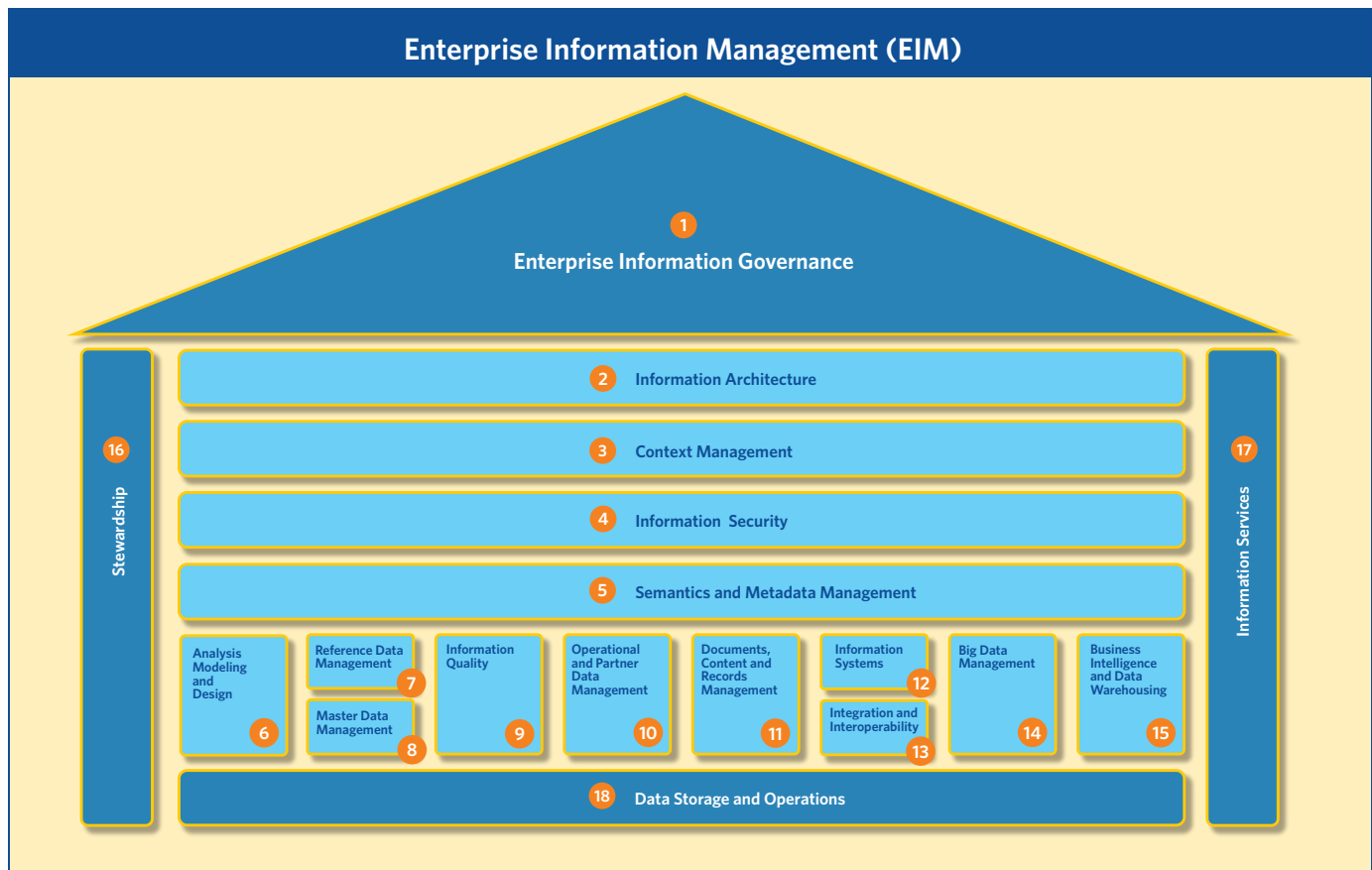


Figure 1. Enterprise Information Management Components and Capabilities

LEGEND FOR FIGURE 1: ENTERPRISE INFORMATION MANAGEMENT (EIM)

1. *Enterprise Information Governance — Guidance, Oversight, Policy Decisions.*

In TopBraid EDG, “governance asset collection” supports capturing policy decisions, documenting governance processes, guidelines, metrics and best practices.

2. *Information Architecture — Data Organization.*

TopBraid EDG can help users organize all their model and data assets.

3. *Context Management — Business Friendly Views and Viewpoints.*

Manages the separation of information architecture layers that deliver business perspectives from the data architecture physically storing data. TopBraid EDG handles enterprise-designed, information layers, as customized collections, that provide business context implemented by architects to organize data, information, and service layers for business use. This describes the “Information Layer,” that is, the business facing views of data warehouses, federated distributed data, business services and their data exchange formats, and the business centric meta-models from analytics tool vendors. Semantic mappings of meaning and preferred terms make the information consistent across these and other layers. Data Lineage mapping relationships demonstrate how the information was developed from many data sources.

4. *Information Security — Data Classification, Data Access and Protection.*

TopBraid EDG provides a rich set of pre-built properties to capture compliance information including compliance with the information security, and privacy guidelines and regulations all to the finest security grain at the atomic “field” level.

5. *Semantics and Metadata Management — Ontologies, Concepts, Preferred Terms, Descriptions, Definitions, Mappings.*

TopBraid EDG uses standard ontology models to define all metadata and other aspects of its functional behavior, views, etc. These models are user configurable -- they can be enriched or modified as needed.

6. *Analysis Modeling and Design — Business Models, Conceptual Models, Data Models, Data Exchange Format Models, Data Specifications.*

One of the strongest features of TopBraid EDG is that it can capture and connect any type of models, specifications and formats — be they about data or about the broader enterprise landscape such as process models.

7. *Reference Data Management — Permissible Values.*

RDM is a focus of one of EDG’s modular packages. Data elements governed within data asset collections refer to reference data as permissible values.

8. *Master Data Management — Identity, Standardization.*

Depending on the tools you already have, TopBraid EDG can store and govern master data or it can integrate with dedicated master data management systems.

9. *Information Quality — Profiles, Measurement, Metrics.*

TopBraid EDG supports information quality initiatives by providing capabilities for defining business rules and quality metrics, integrating with the data profiling tools and storing outputs of data scanning and profiling.

10. *Operational and Partner Data Management — Formats, Mappings, Contracts.*

TopBraid EDG can capture contracts with partners and mappings between metadata and controlled values used internally to those used by partners and vendors. Once this information is captured, TopBraid EDG can be used to perform the required translations.

11. *Document, Content and Records Management — Document Metadata, Taxonomies, Tags, Indexes.*

TopBraid EDG can be used to govern structured and unstructured data. It includes support for taxonomies, content corpora and content tags. Automated tagging is available through the TopBraid Tagger and Autoclassifier module.

12. *Information Systems — Business Systems, IT Systems, Software Platforms, Software Components.*

Information about these assets is captured using TopBraid EDG technical asset collections.

13. *Integration and Interoperability — Data Exchange, Data Flows, Translations and Transformations.*

The Data Lineage asset collection of TopBraid EDG captures information about data flows and transformations.

14. *Big Data Assets — High Volume, High Velocity, Platforms.*

In TopBraid EDG, big data asset collections are used to govern information not only about the datasets in data lakes, but also the overall big data infrastructure such as nodes, controllers and jobs. EDG can also be used to support movement of data from the operational data stores into data lakes, for example, by automatically generating AVRO schemas and capturing what data can be archived and when this can happen.

15. *Business Intelligence and Data Warehousing — Traditional BI Platforms, Reports, Analytics.*

TopBraid EDG makes analytics more powerful by governing BI hierarchies and assisting with data aggregation and data flows across all data sources.

16. *Stewardship — Asset Level Responsibilities.*

The RACI matrix capability in TopBraid EDG captures responsibilities at the desired level of granularity - either an entire asset collection or an individual asset.

17. *Information Services — Data Services, Business Services, Metadata Services.*

TopBraid EDG offers many pre-built services. It also offers powerful tools to make it easy for users to configure additional services.

18. *Data Storage and Operations — the Data itself, its Data Structure, Data Architecture, Operational Metadata, Measurements, Metrics, Jobs, DevOps.*

TopBraid EDG is able to support governance of all data sources and technology components

As an application, TopBraid EDG has implemented key use cases for information governance and EIM. Its user interface is consistent across all assets and becomes more powerful as a user comes to know it. It is a power tool for an enterprise to know, model, manage, and “implement itself.” (See Sidebar I, below.)

Because TopBraid EDG uses layered models for each type of asset with rich metadata based on standard ontologies, users can ask questions directly against the model, against the data instances, and/or against the model plus instances by using EDG’s powerful search and visualization capabilities, or by using the semantic query standard called SPARQL. Further, advanced stewards can easily extend TopBraid EDG to support their unique requirements by working with the integrated development environment (IDE) provided with it — TopBraid Composer Maestro Edition.

Users can start small and rapidly grow into more advanced capabilities (see **Sidebar I II, page 7**). Enterprise priorities change often and staff need to adapt quickly. TopBraid EDG makes it easy to take a bite at a time, shift direction, stop for a while, add some automation, and not lose a thing. It keeps full history so you can easily follow through later. While auditors will love this, it is the practitioners themselves who will use it regularly as they identify quality issues, fill gaps, fix mistakes, measure effectiveness, and automate upon it.

In the rest of this white paper we will focus on a specific subset of information — reference data, and explore how TopBraid EDG can help enterprises to mature RDM.

SIDEBAR I: Topbraid EDG — An Agile Data Governance Solution



Actionable understanding of enterprise information requires connecting business, technical and operational metadata into a single data landscape. Many data governance solutions, based on proprietary approaches, are limited and can’t effectively meet this goal. Instead, they may simply create more data silos.

TopBraid Enterprise Data Governance™

(TopBraid EDG) is a new type of agile data governance solution. It uses a non-proprietary, graph standards-based, model-driven approach to capturing and preserving the meaning of data — something we call **semantic information management**.*

In contrast to traditional approaches, the semantic approach can capture both business and technical metadata. It is flexible and can adapt to changes in the data, metadata, business needs and the organization itself — **because connections are important!**

* See also our whitepaper: **How Can Semantic Information Management Help to Preserve Meaning in a Dynamic Data Environment?**, available here: topquadrant.com/resources/whitepapers

THE IMPORTANCE OF REFERENCE DATA AND ITS EFFECTIVE MANAGEMENT

Reference data is found in every application used by an enterprise including back-end systems, front-end commerce applications, data exchange formats, and in outsourced, hosted systems, big data platforms, and data warehouses. It can easily be 20 to 50% of the tables in a data store. And the values are used throughout the transactional and mastered data sets to make the system internally consistent. How well it is managed has a major impact on every aspect of an organization's use of data — from the integrity of its business intelligence reports, to the success or failure of its system integration efforts.

Reference Data Management (RDM) is an integral part of overall Enterprise Information Management (EIM). Managing reference data well requires aligning its governance with other key functions of EIM. If done well, this has a high payback. Done poorly, it has a high cost.

In the whitepaper *Maturing Reference Data Management*, we describe a Reference Data Management road map using a maturity model, with the five levels shown in Figure 2, (see below), as one likely path an enterprise can take to mature RDM, and the benefits it will gain as a result.

In the whitepaper *The Foundations of Successful Reference Data Management*, Malcolm Chisholm discussed the challenges associated with implementing a reference data management solution, and the essential components of any vision for the governance and management of reference data. Malcolm addressed the following topics: What is reference data? Why is reference data management important? What are the challenges of reference data management? What are some best practices for the governance and management of reference data? What capabilities should you look for in a reference data solution?

Both of these and other whitepapers are available for download from TopQuadrant's website:

topquadrant.com/resources/whitepapers

In this whitepaper, we further discuss connections between RDM and EIM and then describe how TopBraid EDG can help enterprises move to the higher levels of RDM maturity and, more generally, support best practices for information management.

We have also mapped the best practices and key capabilities for RDM described in *The Foundations of Successful Reference Data Management* to the RDM maturity levels described in the *Maturing Reference Data Management* whitepaper (see Table 1, page 3).

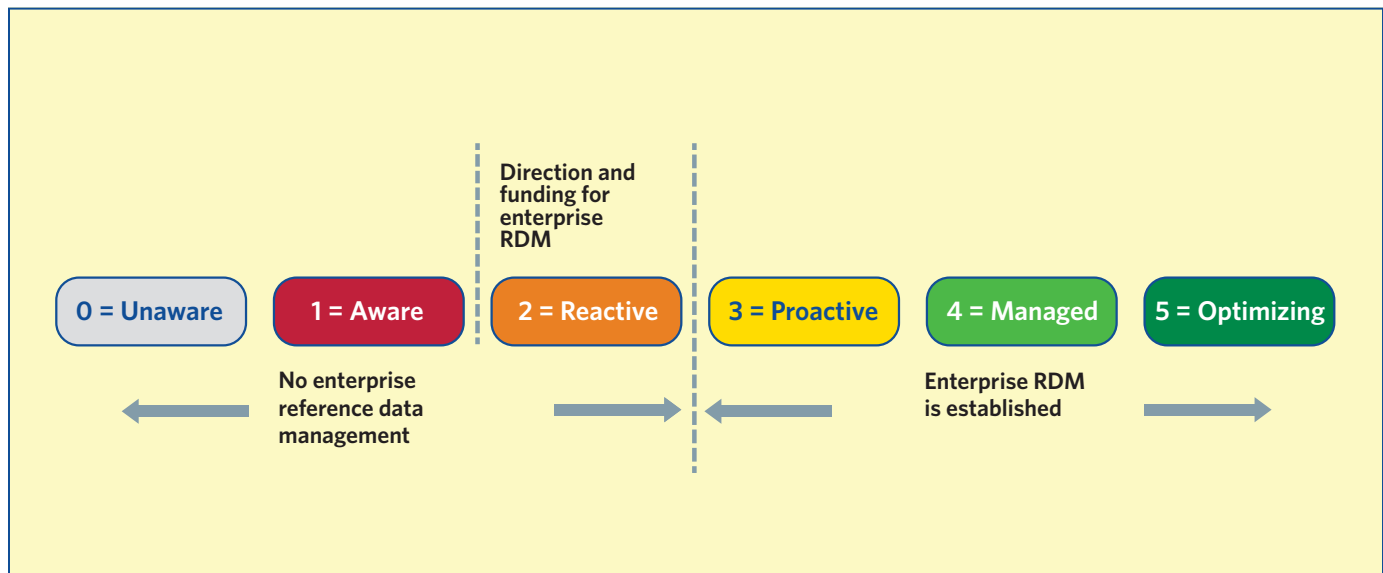


Figure 2. Maturity levels and high level maturation steps

HOW TOPBRAID EDG HELPS WITH REALIZING RDM MATURITY

RDM as Part of EIM Requires Flexible, Comprehensive Support

As an enterprise evolves its RDM, its Enterprise Information Management (EIM) processes coevolve. Business and data analyses move closer together by seeking common terminology and meaning through semantic analysis and mapping. Data modeling shifts up to standards-based, layered models to enable connected abstract conceptual models,

through logical models, and on into physical models of data implementation, data storage and movement — with a clear connection to conceptual business meanings. RDM evolves into governance of controlled business terminologies and lists of meanings within a conceptual reference domain. Enterprise Standard Code sets become a hub of meanings that map to physical implementations of systems' code sets.

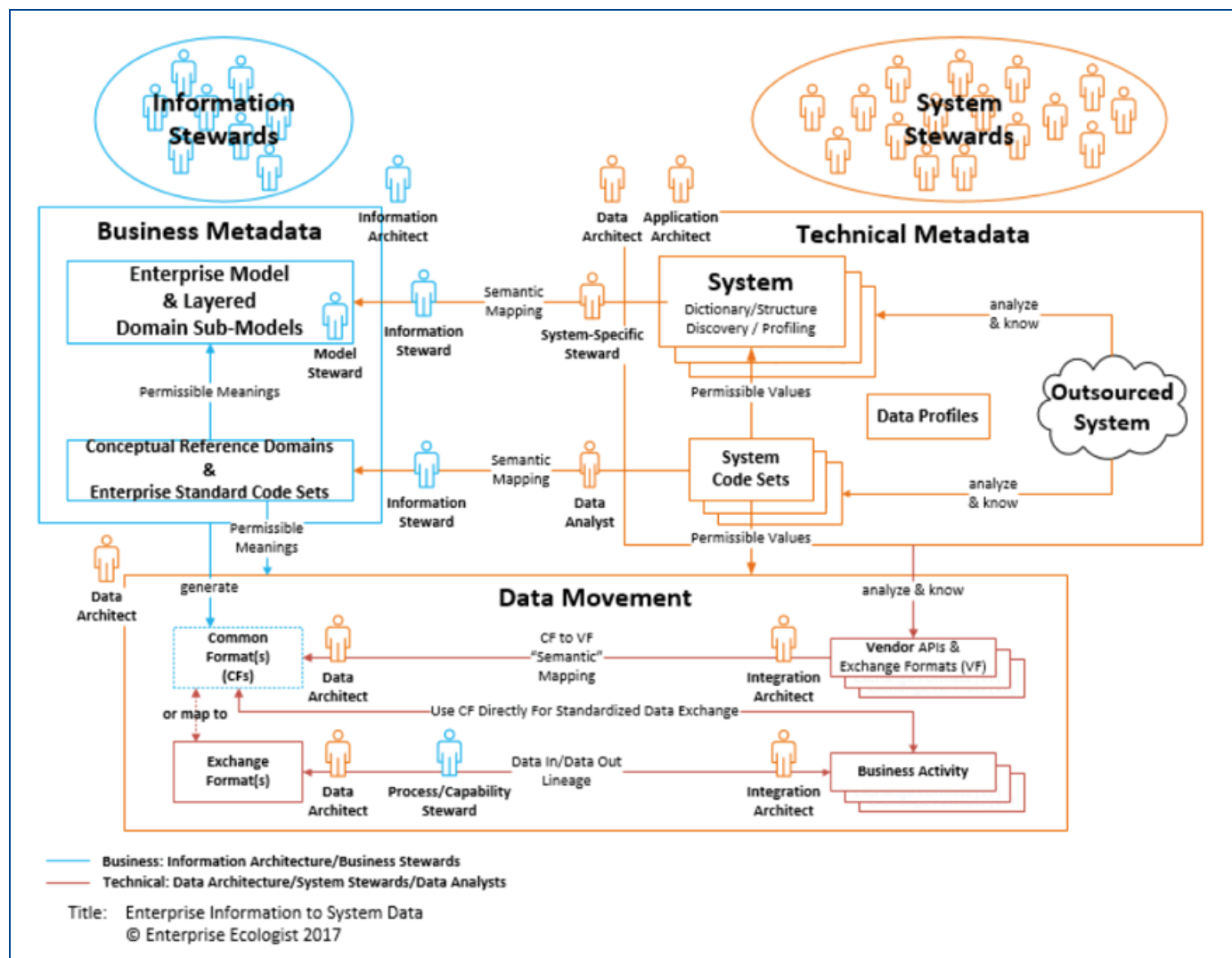


Figure 3. High level EIM assets and roles.

This enables **business semantic integrity** (no loss or confusion of meanings) between people and systems or systems to systems. It allows for more efficient and quicker integrations between systems. It increases business agility by making integration with trading partners faster and easier. By optimizing enterprise's management of its assets, it enables enterprise to effectively execute its business models.

While having a stand-alone focused RDM solution is better than having no solution, moving to the higher maturity levels requires the leverage provided by flexible solutions that can manage inventories of any asset type, and the business and technical relationships between them. Such solutions must have a view into a full enterprise scope of business assets

and technical assets. **Figure 3, (page 6)**, provides a very high level information-centric development lifecycle view of the assets and roles commonly used to implement information systems. These assets have traditionally been captured in a variety of different metadata and configuration management tools.

But at an enterprise level, EIM requires tooling that can support an interconnected inventory of all enterprise capabilities and information assets (e.g., code sets in the case of RDM), their knowledge capture, access and contribution to them by the broad range of stakeholders, and the ability to use automation to make information-intensive processes practical and inexpensive.

SIDEBAR II: Governance Packages Available in TopBraid EDG

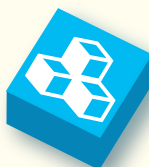
In ramping up a data governance program, different organizations may have differing priorities and starting points. With TopBraid EDG, you can start incrementally. For example, your first focus may be on governing just business glossaries, just reference data, or on metadata management. No matter where you start using EDG, you can always extend your scope to governing other assets when you are ready to do so.

To support this comprehensive but staged approach, TopBraid EDG provides the following packages, any of which are available as an initial configuration of EDG. Each package can also be used in combination with the other packages toward your targeted scope of data governance. Visit topquadrant.com/products/topbraid-edg-gov-packs/ for details on available EDG packages and additional modules.



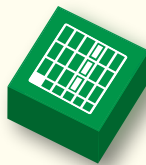
Vocabulary Management

Create, connect, and use taxonomies and ontologies for improving search, clarifying enterprise terminology and enriching unstructured data.



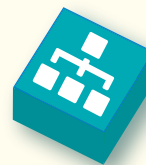
Metadata Management

Govern technical metadata about data assets including databases, datasets, logical, and physical data models. Combine with Business Glossaries to map key data elements to business terms.



Reference Data Management

Profile, govern, update, and provision reference datasets. Use comprehensive metadata to document the meaning of reference data.



Business Glossaries

Define and connect glossary terms or combine with metadata management and establish connections between terms and technical metadata.

TopBraid EDG organizes catalogs by type of asset with complete, built-in metadata to create complete business definition and descriptions, unlimited annotations, semantic mapping between business metadata and technical metadata, the configuration relationships between technical assets, data lineage, and the ability to extend it as needed.

Practitioner-role silos disappear as they collaborate on the information describing the assets and mutual management of the quality of the data itself. The completeness, validity, timeliness/currency, and accuracy ensures the business can trust the information assets to act and automate processes on.

RDM CAPABILITIES IN TOPBRAID EDG

In support of RDM, EDG can model simple to complex code sets including any format (e.g., JSON, XML, csv, tsv, ttl) of reference data coming from external standards bodies, the wide variety of vendor applications, and enterprise developed code sets.

In TopBraid EDG, the code sets are standard, controlled vocabularies within various contexts such as the “enterprise,” “industry,” “system,” “capability,” “ad hoc,” and many other possible contexts. EDG lets users manage these code sets within and across contexts. Establishing semantic relationships from business terms and/or ontologies to other objects ensures that they are well understood in their use and that they move between system boundaries without loss of meaning (e.g., to maintain semantic integrity).

The import capability allows you to map an extract of data against a model for import. In fact, the import can run against different extracts to bring in different data elements, from different sources, as needed, to build (or master) a complete “golden” record of the asset. Because the underlying models are flexible, standards-based ontologies, they can capture relationships between modeled concepts and their data instance values — bridging across different catalogs and data stores, including the custom catalogs enterprises may design.

When information flows across systems, the semantic mapping to each of the system code sets and to the data exchange formats, needs to be based on common meaning and reflect the transformation logic and the resultant change in meaning as it moves through data lineage pathways. A large percentage of such translations and transformations occur on reference data based on the rules defined for using the reference data. They participate heavily in integration logic. One of the highest costs and most fragile parts of enterprise systems — integration — requires semantic analysis, consistency, and understanding of each code set across system boundaries.

While enterprises want to minimize, over time, the number of individual reference datasets, achieving this goal will require analysis and harmonization. The collaborative semantic analysis, modeling, and mapping done during harmonization and the mastering into Enterprise Standard

Code Sets significantly improves reuse, quality, and time to market. Once done, it can be leveraged repeatedly. TopBraid EDG provides pre-built services to translate data coded using one reference dataset into data coded using another reference dataset.

Integrating metadata management with RDM, reference datasets in TopBraid EDG can be used to specify “permissible values” (with the associated “permissible meanings”) possible for certain data elements. Such values are often called “lookups,” “enumerations,” “enumerated values,” “list-of-values,” “data element constraints”; but in use, they are a “code set” that can be well managed enabling data design, data integration, reporting and analytics. These are fundamental information assets. When an enterprise manages these well and views them as a component to be understood, modeled, mapped, moved, translated/transformed, and validated against, all EIM activities become easier. IT activities become easier. The business becomes more agile.

As the quality of reference data gets better and its use more consistent, data practitioners and other stakeholders can take advantage of the enterprise information services such as:

- metadata driven integration
- impact analysis/dependency analysis
- information lookup services
- information directories
- system data dictionaries

TopBraid EDG provides many such services “out of the box” and makes it very easy to create additional services. Like a hologram, as the pixels are rendered, and become denser, the picture becomes clearer and understanding, transparency, and the ability to leverage the assets for the enterprise becomes possible.

Table 1 (page 9), describes how RDM capabilities discussed in the “The Foundations of Successful Reference Data Management” are supported by TopBraid EDG. It also maps the capabilities to the RDM maturity levels discussed in the **Maturing Reference Data Management** whitepaper.

TABLE 1.
Malcolm Chisholm's RDM capabilities and related best practices, RDM maturity and TopBraid EDG support

Capability	Description	Matures in level	TopBraid EDG Support
Ability to create a profile of an external reference data standard.	<p>External reference data is maintained by authorities outside the enterprise. It needs to be discovered, selected and understood before an enterprise decides to use it.</p> <p>This capability is also used to track all interactions with the external authority and to assess its reliability.</p>	3-Proactive	TopBraid EDG offers a flexible approach to cataloging assets. With it, users can capture and describe an external reference data standard. Then, through analysis and discussion, decide whether to use it directly, map it to the reference data currently in use in the enterprise, and so on.
Ability to create a profile of a reference data set maintained by an external authority.	<p>Once external reference data has been set up, it needs to be kept current. Capturing information such as "update frequency" is key to be able to keep up to date with changes and new developments. Subscription management also ensures that changes are detected and assimilated as rapidly as possible.</p>	3-Proactive	Each reference dataset in TopBraid EDG has rich metadata. It includes information about onboarding an external reference dataset, its update frequency, relevant subscription information and more. If built-in metadata fields provided are insufficient, users can easily extend the metadata with their own fields.
Ability to perform semantic analysis of each element in the dataset and identify the business concepts that it maps to.	<p>Metadata is required to describe the dataset and each element in it.</p> <p>Semantic analysis may involve managing decisions by appropriate SMEs and stakeholders about what these mapping decisions are, all of which generates even more reference data metadata that needs to be captured.</p>	3-Proactive	With TopBraid EDG, each element in the dataset (column) exists as a property in an ontology. It can be mapped to glossary terms, technical metadata and other managed assets.
Ability to properly document the semantic analysis after it has been performed.	<p>This may include facts about the reference dataset or individual codes. Such facts help users of the reference data understand how to interpret and use it.</p>	3-Proactive	Each element has description fields, connections to data and application requirements and other metadata. These are specific metadata fields. The "fact" field can be used for more general, "catch all" statements. Facts are supported at both the dataset and individual code level.
Ability to import external or internal reference data into a central repository.	<p>Such import must include capabilities for extraction, filtering, transformation, and enrichment. As much as possible, this should be metadata driven.</p>	3-Proactive	TopBraid EDG offers rich import capabilities that are highly flexible and user configurable. Any transformation can be included as part of an import.
Establish and Manage Enterprise Standard Code sets	<p>Begins harmonization across systems and standards with strategic choice of what to work on based on contextual scope, usefulness, project priority, and other factors.</p>	3-Proactive	TopBraid EDG lets users identify some reference datasets as "Enterprise Standard Code sets" for a given entity e.g., country. Users can then map between the code set used by each specific system and "Enterprise Standard Code sets" by creating crosswalks.

TABLE 1. (continued)

Capability	Description	Matures in level	TopBraid EDG Support
Ability to assign accountabilities for all aspects of reference data management per reference dataset, particularly for internal reference datasets.	<p>Internal reference data is for business concepts that are completely specific to the enterprise. It requires a federated approach, because it is created and managed by many different subject matter experts (SMEs). The central RDU must ensure that groups accountable for internal reference data use a standardized approach.</p> <p>This achieves the federated governance model needed for internal reference data. Obviously, this capability requires a rich set of metadata elements for reference data.</p>	4-Managed	<p>TopBraid EDG lets you define a RACI matrix. RACI can be specified for an asset collection such as a reference dataset or, alternatively for an asset such as an individual code.</p> <p>Task assignments and configurable, targeted notifications support collaboration across the enterprise.</p>
Ability to track changes to reference data	For example, if an external reference dataset changes or new values are added into operational systems.	4-Managed (in catalog) 5-Optimizing (in data use)	All changes are audit trailed in TopBraid EDG. Users can see history of changes and, if desired, revert them.
Profile Reference Data Quality in Usage	Profiles actual data where code values are used to categorize it to see if new, incorrect, invalid, or other non-permissible values are in actual use. Errors result in Data Quality triage to disposition activities which may include a cycle of analysis and management in RDM.	5-Optimizing	<p>Since many applications store reference data locally, invalid use of reference data is typically due to the local reference data being out of date with the enterprise standard. TopBraid EDG offers pre-built services for verifying local reference data against the code sets governed in EDG. Results of such periodic verifications are stored over time letting users see any issues with the quality of local reference data and how these issues have been resolved.</p>
Ability to distribute reference data.	<p>Reference data is used widely throughout the enterprise. It is vital that all applications have synchronized copies, so distribution must be addressed. This requires a variety of approaches ranging from the fully automated to the fully manual. However, these approaches must be chosen carefully to maintain operational efficiency.</p> <p>A variety of distribution mechanisms such as exports, web services and ESB integration should be provided.</p>	5-Optimizing	Most capabilities of TopBraid EDG that are available through its user interfaces are also available as RESTful web services. This includes data exports. Reference data can be provisioned to applications as a service that exports an entire reference dataset, subset of it or information about a single code. Services can be scheduled or accessed on demand. There is also a pre-built integration with Enterprise Service Buses (ESB).

SIDEBAR III: A Practitioner's View — From David Chasteen

As an enterprise practitioner, not a consultant, I care about not just any slice of EIM, but all slices, highly interrelated and operating at the same time. I want to be able to automate on top of it and provide information services easily. I want to be able to work on business metadata of terms, vocabularies, and conceptual models. I also want to be able to represent data in any of its forms and the ways it is organized, moved, and manipulated. And I want my co-practitioners to also be able to do it collaboratively — so that we don't create silos and, collectively, we can manage our information assets without having to recreate the wheel.

Here are a few of my other "wants:"

- I want to be able to start small with an incremental bit of knowledge and, then move to another bit to put out a fire, and later, to slowly build out catalogs and through relationships, enterprise level information assets
- I want to work on the assets I need to work on, set them aside for other priorities, and not lose a thing. In other words, I want to be able to change the tires on the car while it is racing to its next destination, and change direction quickly when needed
- I want my investment in information assets to increase in value as I add more and more assets. I want it to increase in value as it gets used more and more
- I want my tools, knowledge, and skills to develop, become sharper, and more able to do what I need them to do

Vendors seldom support the richness, openness and extensibility needed to meet these requirements. But in the case of open, semantic standards-based tools, there exists a platform where I can build the knowledge assets to do just that. I believe TopBraid EDG provides the set of capabilities to do this.

One good place to start is by using the TopBraid EDG Reference Data Management package. You can start

working on a vertical domain of knowledge, for example geography, and adopt open, standard geographic data, extend it with enterprise geographic boundaries such as sales regions, service regions, competitor boundaries and capture the knowledge about it.

If this information exists in other systems, I want to be able to federate connections and mappings to them and extend their functionality by:

- adding meaning
- enriching code values and hierarchies
- capturing the ownership of assets and responsibilities around them
- leveraging any information, in context, that is needed to know, understand, and use the asset directly in my systems

And I want the same assets to make it possible to consistently get answers to my business questions. TopBraid EDG has captured many of the ways a practitioner may use or manage reference data. It manages the tasks such as:

- defining a code set
- modeling the structure of a code set
- managing any rule factors for a code set mapping
- importing the permissible values for a code set, and so on

It even adds a layer to help manage the vendors and versions of code sets coming from external vendors such as standards bodies and standard code sets (e.g., reference data). All-in-all, TopBraid EDG and its RDM, Vocabulary Management and other data governance capabilities provide much of what I need as an enterprise practitioner.

Evaluating Topbraid EDG is Easy and Free

To learn more about TopBraid EDG, visit:

- topquadrant.com/products/topbraid-edg
- topquadrant.com/products/topbraid-edg-gov-packs

Videos and webinar recordings about TopBraid EDG can be accessed here:

- topquadrant.com/knowledge-assets/videos

Contact us at edg-info@topquadrant.com to:

- Discuss data governance your requirements
- Request a more targeted demo of TopBraid EDG
- Ask for a free TopBraid EDG evaluation account



ABOUT THE AUTHORS



David Chasteen **Enterprise Ecologist**

David Chasteen has worked in information-centric roles for over 30 years, primarily within large enterprises (e.g., health care, government, software companies, consulting companies, and as a scientist). Unlike a consultant or contractor, David has had long-term enterprise careers. As a result, he has been able to experience, grow, and iterate on Enterprise Information Management practices from conception to operation. David has led enterprise information modeling, metadata and reference data management practices and been an active participant in most other information-oriented roles and practices.

In his long career, David has wrestled through the aftermath of successful and failed projects, internal politics, consulting engagements,

contractor losses, internal reorganizations, management fads, IT fads, the (wrong) belief that staff are fungible/ interchangeable, bad and good management, no funding, too much funding/no staff, bait and switch consulting, inadequate vendor contracts, knowledge loss from turnover, waves of retirements, and cloud-sourcing. He is proud of being able to keep the efforts alive through them all. Like other dedicated staff, he cared enough and maintained the fundamental belief that information assets are not optional, they are foundational. And that developing deep and wide skill in the various practices pays great dividends.

Today, David has retired into a writing and mentoring life. He is looking forward to sharing what he knows with other passionate practitioners.



Irene Polikoff **Co-founder / CEO, TopQuadrant**

Irene Polikoff has more than two decades of experience in software development, management, consulting and strategic planning. Since cofounding TopQuadrant in 2001, Irene has been involved in many information management projects across a broad range of customer organizations in different industries. She has written strategy papers, trained customers on the use of the semantic technology standards, developed ontology models, designed solution architectures, and defined best practices. With the introduction of TopQuadrant's Semantic Solution Platform and TopBraid Suite, Irene has been responsible for TopQuadrant's product strategy and its data governance vision.

Before starting TopQuadrant, Irene was a principal and a senior development manager at IBM Global Services. Prior to that, Irene held IT management positions at Fortune 500 companies where she was responsible for development and deployment of enterprise-wide mission critical information systems.

Irene is a coauthor of a book on software requirements and architecture — *Capability Cases: Solution Envisioning Approach*, author of many published articles and whitepapers as well as a frequent speaker at conferences. Irene has been actively involved in technology standardization. She currently serves as a co-chair of the W3C SHACL working group that has developed a standard for describing and validating RDF data.

CONCLUSION

This paper provides a brief overview of the importance of information governance — and data governance as a key component of it. It outlines Enterprise Information Management (EIM) capabilities as a foundation to align data governance practices to in order to mature and perform them well. TopBraid Enterprise Data Governance (EDG) is described in terms of how it supports EIM capabilities, based on semantic standard-based approach that enables EDG to manage the entire range of enterprise information assets and their connections.

The paper then focuses on Reference Data Management (RDM) as an integral part of EIM. It discusses how TopBraid EDG can help with progressing “RDM maturity” as it is presented in terms of formal RDM practices in the companion white paper to this one: *Maturing Reference Data Management*, (topquadrant.com/solutions/maturing-reference-data-management/).

About TopQuadrant

TopQuadrant helps organizations succeed in data governance. Its flagship product, TopBraid EDG, delivers easy and meaningful access for all data stakeholders to enterprise metadata, business terms, reference data, data and application catalogs, data lineage, requirements, policies, and processes.

TopQuadrant’s customer list includes organizations in financial services, pharma, healthcare, digital media, government and other sectors.



TopBraid Enterprise Data Governance™

TopBraid Enterprise Data Governance™

is an agile data governance solution for today’s dynamic enterprises. A modular, extendable, standards-based solution, it provides integrated data governance across the ever-growing types of assets and governance needs.

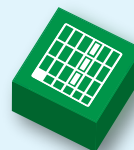
Governance Packages Available in TopBraid EDG



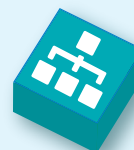
Vocabulary
Management



Metadata
Management



Reference Data
Management



Business
Glossaries

In ramping up a data governance program, different organizations may have different starting points. With TopBraid EDG, you can start

incrementally and add capabilities as you go. For details on available EDG packages and additional modules visit topquadrant.com/products/topbraid-edg-gov-packs/