Managing Metadata to **Leverage Data Governance and Sourcing**

Case Study at a Financial Services Firm

Using TopBraid EVN, this financial services firm provided the business with better and timelier access to information about what data is being used by which applications, along with metrics and reporting on data quality, data lineage, and data connections. To quote one of the firm's project managers: "Business users can now find the information they need without a Ph.D. in SQL."

PRODUCTS USED

TopBraid Enterprise Vocabulary Net (EVN) to connect and govern business and technical metadata across data sources and information domains. With TopBraid EVN, business vocabularies become an integral part of enterprise-level information management.

TopBraid Live (TBL) to dynamically deliver metadata services needed for data transformation, data integration and search.

BENEFITS

• Improved information access by making the knowledge about the meaning of data available to search and data integration tools

• Governance of enterprise information assets

• Improved information quality across the organization

• Ability to facilitate seamless flow of data from one system to another

• Ability to meet regulatory compliance requirements

• Lower costs through minimizing redundancy of data

• Ability to flexibly model any metadata and data using W3C standards

• Integration with other enterprise applications through REST interfaces and other enterprise service buses

Background on the Financial Services Firm

The company is a large North American financial services firm that transacts in the capital markets. It has an international reputation for innovation and leadership in operations.

Challenges

With the business goal of building capabilities to support more effective decision-making, the company was looking to build a metadata repository to accomplish several objectives:

• Leverage data sourcing and governance work

• Quantitatively measure process improvements for data governance

• Make data and business analysis sourcing and processing more efficient

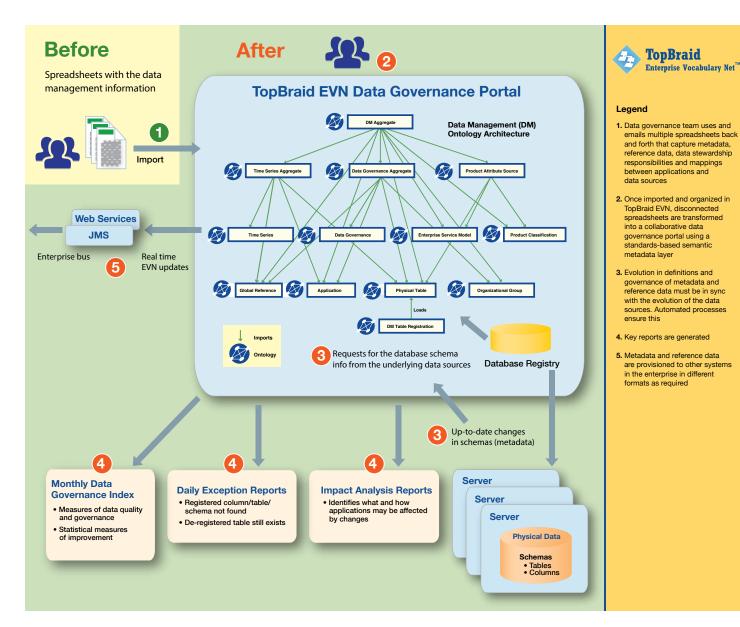• Generate metrics on data usage and consistency

For the firm to accomplish these goals, it required techniques to capture diverse metadata. The company decided to use the World Wide Web Consortium's (W3C) semantic web standards because of the flexibility they provide for working with heterogeneous data and to take advantage of the growing market adoption of the standards-based technologies. With the capabilities of TopQuadrant's Enterprise Vocabulary Net (TopBraid EVN), the company was able to effectively support this initiative.

TopQuadrant™

TopQuadrant's standards-based solutions enable a semantic ecosystem among people, applications and data—lowering the cost of ownership and enabling intelligent, data-driven action. TopBraid Enterprise Vocabulary Net (TopBraid EVN) is a flexible, web-based system for managing semantic information models. It enables business stakeholders to collaborate on defining and linking enterprise vocabularies, taxonomies, thesauri and ontologies used for information integration, navigation and search.

For more information visit www.topquadrant.com, or contact us at evn-info@topquadrant.com or by phone at +1 919 300 7945.

Ask us about scheduling a demo to explore how TopBraid EVN meets your specific requirements.

Before

Spreadsheets with the data management information

1 Import

After

2

TopBraid EVN Data Governance Portal

Data Management (DM) Ontology Architecture

DM Aggregate

Time Series Aggregate | Data Governance Aggregate | Product Attribute Source

Time Series | Data Governance | Enterprise Service Model | Product Classification

Global Reference | Application | Physical Table | Organizational Group

Loads

DM Table Registration

Imports
Ontology

3 Requests for the database schema info from the underlying data sources

Database Registry

3 Up-to-date changes in schemas (metadata)

Web Services
JMS

Enterprise bus   Real time EVN updates

5

4 **Monthly Data Governance Index**
- Measures of data quality and governance
- Statistical measures of improvement

4 **Daily Exception Reports**
- Registered column/table/schema not found
- De-registered table still exists

4 **Impact Analysis Reports**
- Identifies what and how applications may be affected by changes

Server
Server
Server
Physical Data
Schemas
- Tables
- Columns

TopBraid
Enterprise Vocabulary Net™

Legend

1. Data governance team uses and emails multiple spreadsheets back and forth that capture metadata, reference data, data stewardship responsibilities and mappings between applications and data sources

2. Once imported and organized in TopBraid EVN, disconnected spreadsheets are transformed into a collaborative data governance portal using a standards-based semantic metadata layer

3. Evolution in definitions and governance of metadata and reference data must be in sync with the evolution of the data sources. Automated processes ensure this

4. Key reports are generated

5. Metadata and reference data are provisioned to other systems in the enterprise in different formats as required

Solution

The first step for the firm was to connect multiple data types such as databases, spreadsheets and documents that were stored in a content management system, in XML and other formats and put them into a common framework. Since data definitions change, the firm's requirements for the metadata management system included managing and updating of changes from each data source and the retaining of a full audit trail of all changes.

Once the data had been connected, name-alignment transformations were defined to assign Uniform Resource Identifiers (URIs) to all enterprise concepts, so that the data could be shared, merged, and re-used across multiple applications. This allowed business users to locate different types of data across formats, applications and subject areas. Covered subject areas include the following:

- Countries
- Currencies
- Financial Instruments
- Legal Entities
- Products and Assets
- Time Series Data

TopBraid EVN is an open and highly configurable system easily extended with customer-specific rules, reports and user interface customizations. To satisfy the core requirement of providing real-time compiled information for quarterly reporting, TopBraid EVN was extended to produce reports for the firm's Data Governance Index (DGI). DGI is a quantitative metric designed by the company to measure the degree of governance of its Enterprise Data. An internal council uses this to measure data governance improvements. DGI allows the firm to properly identify how data is being used in various divisions throughout the organization.

Another core requirement was to develop a curated table registry of physical data assets used in the enterprise. This registry defines data assets that the data management team tracks and enriches with metadata from various sources, all integrated by TopBraid EVN. Daily exception reports are produced to register new database tables and to track changes in physical tables to ensure real-time accuracy of data management measures.

Results

Currently, the firm has captured more than 13,000 types of data across multiple interconnected vocabularies. Using the DGI index, the company ensures that its data has managed processes and controls, identified data stewards, required quality checks, service level agreements, and other data quality measures.