Case Study:

Harnessing Semantic Metadata to **Enhance Content Access**

Mayo Clinic uses TopBraid Enterprise Data Governance–Vocabulary Management (EDG–VM) to deliver intelligent results to end users

"*The actionable information provided to our end users created a more seamless search experience. Using this system gave us the freedom to create relationships between our existing vocabularies and helped us transition our existing websites into one primary source of information, reducing confusion and providing users better access to information they are seeking through improved access to relevant content and enhanced search results.*"

— Mayo Clinic

PRODUCTS USED

TopBraid Enterprise Data Governance–Vocabulary Management (TopBraid EDG–VM) — a flexible, web-based solution for managing business vocabularies as semantic information models. TopBraid EDG–VM was designed to respond to the needs of organizations to get more value out of the enterprise's data by using business vocabularies and metadata to enhance search, navigation and content delivery, align the meaning of data across data sources and integrate diverse information.

TopBraid Live—an intelligent business applications server and the cornerstone of the TopBraid architecture. Offering all capabilities of the TopBraid platform, it leverages semantic technology to help customers connect silos of data, systems and infrastructure and to build flexible applications from linked data models. All TopBraid server products use TopBraid Live as their foundation.

BENEFITS

Through focused explorations of available technologies, Mayo Clinic determined that the capabilities of TopBraid EDG–VM and TopBraid Live would help to meet the goals of their Knowledge Content Management System, including to:

• Integrate three external websites into one single source for trustworthy medical information at *mayoclinic.org*.

• Increase web search efficiency and accessibility.

• Increase the time viewers spend on the website.

• Automatically generate new content using linked vocabularies.

Background on Mayo Clinic

A worldwide leader in nonprofit medical care, Mayo Clinic is dedicated to providing a wealth of medical information applicable to all walks of life. Mayo Clinic has a rich history of leadership in sharing medical information and was among the early innovators of providing extensive health information on the Internet. In January 2014, Mayo launched the next evolution of its presence on the Web by integrating its three existing external websites—*mayoclinic.com, mayo.edu* and *mayoclinic.org*— into a single platform.

This milestone is the outcome of Mayo's Knowledge Content Management System (KCMS) initiative—the single largest project undertaken in the history of Mayo Clinic's Web assets. KCMS is a large-scale solution that combines traditional content management capabilities with the power of semantic technology to integrate healthcare knowledge for the enterprise. Now, the millions of people who use Mayo's websites to find answers to their medical and lifestyle questions get improved access to relevant content and enhanced search results at the semantically enhanced *mayoclinic.org*.

Challenges

For more than 100 years, Mayo Clinic has delivered high-quality healthcare by integrating knowledge across patient care, research and education. In the digital age, Mayo found that its different departments and practices were growing increasingly disparate with isolated silos of knowledge. Updating and re-invigorating its integrated approach to knowledge sharing in the digital environment required an innovative solution.

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. TopBraid Enterprise Data Governance–Vocabulary Management (EDG–VM) supports business stakeholders who need to collaborate on defining and linking enterprise taxonomies, ontologies and thesauri. They use these to improve search, enhance content navigation and align the meaning of data across data sources.

Visit www.topquadrant.com, or contact us at edg-info@topquadrant.com or by phone at +1 919 300 7945.

Ask us about scheduling a demo to explore how TopBraid EDG–VM meets your specific requirements.

Submits Search → **Search Application**

Business User

1 Submits Search

5 Adds value to user's search by enabling navigation to contextually related information using the business vocabularies.

3 Adds value by semantically rewriting queries.

Query **2** **6** Results

Query Engine
Keyword Search | Semantic Search

TopBraid Live™

Business Vocabularies (taxonomies / ontologies) **1**

Business Vocabularies (taxonomies / ontologies)

Tag Set

4 Index

TopBraid Enterprise Data Governance™

Vocabulary Management

4 **TopBraid Tagger™**

Search Indexing Engine (FAST, Lucene, SOLR, …)

TopBraid EVN's AutoClassifier Engine

Adds value by providing curated vocabularies to be used by TBL terminology server at run time to enhance user's search.

Adds value by creating and managing vocabularies needed by the text extraction engine.

2 **3**

Adds value by enabling editorial curation of autogenerated tags and direct manual tagging.

Content

Sharepoint Sites | Documentum, OpenText, and other CMS | XML | Network Folders | Email | …

Run-time actions / tasks *(numbers in orange)*

1. Search User submits a request
2. Search application sends it on to the query engine
3. Query engine uses TopBraid Live (TBL) to re-write the query
4. Query is then executed and results returned
5. Prior to returning results to the user, TBL adds related links using the business vocabularies
6. Search application incorporates the enhanced search results into resulting display

Back-end tasks / pre-search tasks *(numbers in green)*

1. Using TopBraid Enterprise Data Governance Vocabulary Management (TopBraid EDG–VM), taxonomists develop vocabularies
2. Indexer (part of the search engine), crawls the documents and develops the index
3. Auto-tagging engine may use the vocabularies to extract information and enrich the index
4. Using TopBraid EVN Tagger, content owners may do manual tagging and also review and refine automatically generated tags

TopQuadrant™

Figure 1: TopBraid EDG–VM and TopBraid Live provide semantic services within an overall search/query infrastructure to provide users enhanced access to digital content.

Mayo did extensive research and testing of available content management systems (CMS) and determined that these systems do not meet their needs for management of terminologies and metadata. In order to successfully achieve their goals, Mayo Clinic required a solution that extends traditional content management technologies with a *semantic services layer* to support standards-based knowledge interoperability within and between its Web properties.

Solution

In response to these challenges, Mayo has been deploying KCMS to multiple stakeholders in well-designed stages. The solution was first deployed to Mayo's 'Ask Mayo Expert' facility, and then to its Research Web. Following these successes, Mayo deployed KCMS to *mayoclinic.org*, its large scale public website that provides medical information to consumers world-wide. Beyond this, KCMS will be deployed to clinical practitioners. Though the specific capabilities needed may differ somewhat across these deployments, the overarching goal for all stakeholders from researchers and clinicians to end users is improved access to relevant content and enhanced search results for users.

At its core, the KCMS solution implements a *semantic services layer* using TopBraid Enterprise Data Governance–Vocabulary Management (TopBraid EDG–VM) and TopBraid Live (TBL) to provide native semantic web standards-based capabilities for content enrichment, search enhancement, taxonomy management and run-time terminology services. Combined, these capabilities enable users to quickly find the right information, regardless of the author or the storage method. See Figure 1 for an illustration of the capabilities of TopBraid EDG–VM and TBL. *(Note: this figure does not reflect the specific architecture of the KCMS solution at Mayo, but illustrates similar capabilities of TopBraid EDG–VM and TBL that are being used within KCMS.)*

TopQuadrant's Bob DuCharme, Director of Digital Media and Solution Architect, who worked with Mayo on this project, observed: "The project involved the migration of several large-scale websites using separate infrastructures to a single unified infrastructure, re-using as much metadata as possible from old infrastructures about topics and subjects covered by each page. Existing CMS packages had no way to migrate that data because of its ad hoc nature, but automated processes let us convert the data to take advantage of standardized semantic web models, SPARQL queries, and the TopBraid tools that build on these to tailor the data according to the needs of the new system."

Results

The solution allowed for the dynamic creation of alternate links and the generation of new sections on Mayo Clinic's website such as lifestyle guides and doctor index lists, which are populated with relevant content automatically based on the relationships formed between vocabularies.

Employing TopBraid EDG–VM's flexibility to manage semantic vocabularies and create semantic metadata for tagging content allowed Mayo Clinic to create relationships among vocabularies and between terms and content. This provided a more intelligent means of searching for and delivering results on Mayo Clinic's integrated website. In turn, this helped make the experience of visiting and searching on Mayo Clinic's site more user-friendly and adaptable.