



SAS® Ontology Management

Maximize the value of your text repositories by linking them together with consistently and systematically defined relationships

What does SAS® Ontology Management do?

SAS Ontology Management, powered by Teragram's advanced linguistic technologies, creates semantic terms that are used to organize previously disassociated and isolated text repositories. It enables you to create relationships among key information so subject-matter expertise can be systematically built into information search-and-retrieval activities. Domain experts define the rules of relationships among key text information concepts (within repositories and across taxonomies and document collections), specifying an organizational thesaurus that automatically links domain knowledge across the enterprise.

Why is SAS® Ontology Management important?

It helps organizations maximize the value of their semantic repositories and collections by linking them together with consistently defined relationships to quickly and accurately address information search-and-retrieval needs.

For whom is SAS® Ontology Management designed?

This product is designed for organizations that are beginning to address their enterprise textual data strategy. It also is designed for organizations with existing and isolated repositories of thesauri and vocabulary collections and/or pre-existing and siloed taxonomies that require a cohesive strategy so that meaningful, related information can be returned to complex queries.

With the escalating growth of information in an increasingly globalized society, there is a corresponding need for subject-matter expertise. Simply returning data without relating the meaningful relationships that define knowledge is no longer sufficient. Subject-matter expertise requires that pieces of apparently unrelated data be understood in relationship to each other.

SAS Ontology Management, powered by Teragram technology, creates and maintains consistent and centralized metadata across textual repositories so that information search-and-retrieval engines systematically identify common concepts. The result is that complex questions that may not even have the answer explicitly located in the text can be satisfied with meaningful responses.

Key Benefits

- Extend the reach of knowledge workers by automating their subject-matter expertise.** With SAS Ontology Management, subject-matter knowledge can be encapsulated in ontology definitions. Users who define the ontologies are subject-matter experts who understand the associations and relevance of topical concepts. Once defined, ontologies are executed during the processing of input documents, freeing knowledge workers from evaluating every incoming bit of text.
- Integrate text collections across the enterprise for consistent understanding and improved search-and-query retrieval.** Using semantic terms and attributes to identify key interrelationships, you can relate
- seemingly unrelated documents and content repositories. The GUI interface of SAS Ontology Management makes it easy for knowledge workers to identify semantic terms, which are then used to automatically discern whether specific concepts are related to one another. This ensures that text materials are consistently classified for appropriate distribution, and the integrated information provides better answers to complex queries.
- Improve the administration and auditing of enterprise information.** Enterprise administration capabilities define access, authorizations and tracking of changes coming in from different knowledge experts and users defining the ontologies. This maintains consistency in the metadata of definitions. These metadata definitions ensure that text collections and repositories are associated with one another systematically and automatically to provide well-defined, organized and traceable relationships across large document collections.
- Derive meaning and maximum value from enterprise document collections.** SAS Ontology Management integrates existing text repositories and assets by associating disparate collections with common definitions and rules of hierarchy. In addition, it applies a defined structure (ontology) to all new input materials for consistent subject-matter relevance. This enables organizations to maximize the value of their text repositories by linking them together with consistently defined relationships, which helps to quickly and accurately address information retrieval needs.



Product Overview

SAS Ontology Management is a robust solution that enables the management of content and maintenance of ontologies in enterprise content repositories and databases. When integrated with taxonomy management, knowledge workers can maintain metadata across repositories and databases, and automatically tag documents according to the defined taxonomies.

Enables collaborative ontology development and maintenance

An intuitive Windows user interface provides a visual ontology structure that simplifies collaborative ontology project development. It provides visual representations of the complex interrelationships between the various types of information, metadata and instantiations

within projects. Interface components provide important information on project development. Users can assemble complex projects that effectively locate and classify the key information that differentiates subject-matter expertise from knowledge. Centralized administration enables several users to log into and make changes to a database. An administrator approves or rejects the changes so information is updated as it is generated by those working in the field.

Integrates existing document repository assets

SAS Ontology Management integrates silos of text repositories, document collections and multiple taxonomies by associating disparate document repositories and collections with common definitions and rules of association. It

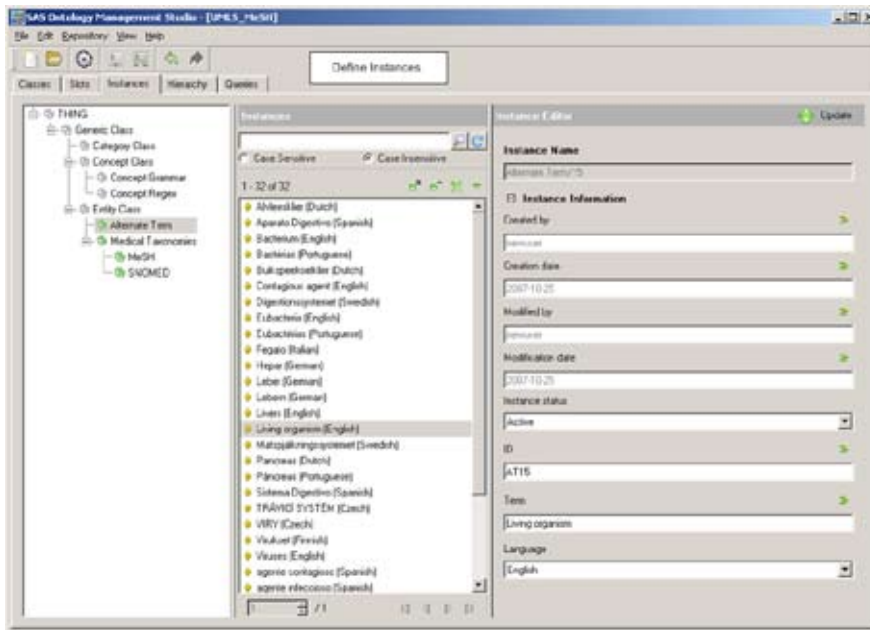
applies a defined structure (i.e., ontology) to all input materials for consistency, and ensures that text repository materials are consistently classified for appropriate distribution. Wizards enable easy importing and exporting of projects and attributes.

Identifies relationships between document repositories

Identifying relationships between seemingly unrelated documents and content repositories provides better answers to complex questions. Semantic terms and their attributes are used to identify key interrelationships among entities, concepts and categories so that subject matter can be defined as either related or not. SAS Ontology Management provides an easy-to-use interface for defining these relationships (i.e., semantic terms and their attributes), and provides a systematic approach to organizing the various types of information that will be accessed and under what conditions. Information stored in a SAS Ontology Management database can be exported to SAS Enterprise Content Categorization to build complex taxonomies that can be used to automatically categorize documents and to extract related entities and events from documents.

Build subject-matter expertise into search-and-retrieval activities

SAS Ontology Management creates semantic terms that are used to organize and relate previously disassociated and isolated bits of information across document collections. It provides the ability to create relationships among key information (semantic terms) so that subject-matter expertise is automatically built into search-and-retrieval activities. Once defined in SAS Ontology Management, these relationships ensure that key information is consistently identified and a uniform approach for surfacing content is created.



SAS Ontology Management provides an easy-to-use interface that simplifies the development and maintenance of ontologies. One of the first steps in the process is to define relationships among semantic terms and their relevance, ensuring that information is consistently identified and uniform associations of content occur.

Consistently applies subject-matter expertise across document repositories in real time

SAS Ontology Management includes a predetermined list of semantic terms to help organizations begin their ontology development. From an easy-to-use GUI, users can add to this list, customizing the relationships among terms to capture subject-matter expertise. The instances of occurrence can be specified for consistent identification of relationships among topics of interest. The resulting hierarchies then can be used as the input taxonomies for automatic content categorization using SAS Enterprise Content Categorization. SAS Ontology Management can also associate content defined in the ontologies at run time, creating metadata when the ontology is connected to a database using APIs.

Centralizes administration for collaborative ontology development

Managing large amounts of information often requires teams of individuals and groups. It is not always necessary or prudent to allow all of the members of your organization or development team equal access to all of the levels of ontology development. For example, it might be necessary for one development group to be able to read the definitions created for some information classifications so they can identify the representations that should be used as part of the classification definitions they are developing. SAS Ontology Management enables the project manager to set default permissions at each of the three project levels, and to override the permissions as necessary for specific users and groups. This permission setting enables all of the development teams or individuals to access different aspects of project development according to their needs and expertise, while also protecting the integrity of the project.

Key Features

Enables collaborative ontology development and management

- An intuitive Windows user interface provides a visual ontology structure that simplifies project development.
- Define ontologies in hierarchies of classes, and subclasses, of semantic terms to help organize the different types of information that will be accessed.
- Parent-child relationships between multiple instances associated with the same attribute (aka “slot”) can be defined.
- “Slots” allow the definition of metadata that forms the definition of classes (their relevance, instantiations, etc.). Standard types include:
 - o Integer (numeric).
 - o Floating.
 - o String (character).
 - o Boolean.
 - o Instance of.
 - o Class ID.
 - o “Slot” ID
- “Slot” constraints can include:
 - o Minimum and maximum values.
 - o Default and candidate values.
 - o Validation against regular expressions.
 - o Uniqueness.
 - o Multiple cardinality.
 - o Allowable classes in cases of instance of, class ID and “slot” ID.
- Inverse attributes (“slots”) can be defined to indicate reverse relationships.
- Instance-level and class-level versioning enable you to revert to previous definitions.
- Automatic verification of data is provided with defined relationships and inverted relationship definitions among parent-child relationships.
- Provides high scalability for users, classes and terms, and multiple ontologies.

Integrates existing document repository assets

- Upload wizards enable import and export of projects and attributes (i.e., instances).
- Nontab characters are available for separating values in multiple cardinality lists to facilitate the import and export of Excel spreadsheets, CSV files, etc.
- Built-in integration with SAS Enterprise Content Categorization is included for the automatic categorization of documents.
- Accepts XML input and creates XML output so if there are taxonomies or lists that exist or are acquired, they can be easily included as a simple import.
- Readily maps to RDF (Resource Descriptive Framework) and OWL formats.
- Output to metadata repositories, including SAS Metadata Server, SharePoint, FAST, EMC Documentum, Endeca and others, is connected via APIs.
- Data can be queried and/or modified via API calls to the server for execution of ontology applications.

Identifies relationships between document repositories

- Provides an easy-to-use interface to define relationships (i.e., semantic terms and their attributes) systematically and comprehensively.
- Organizes rules of relationships between semantic terms into classes and subclasses, which have attribute definitions that define their relevance under specific conditions. Relevance conditions (otherwise known as instantiations) include definitions about the associations, characteristics and constraints specified by the user.
- Characteristics and constraints comprise the metadata relationships and form the definition and operation of the class.
- Rules associate text documents and repositories by their class definitions; not just the key words contained in the materials.

SAS® Ontology Management Technical Requirements

SAS Ontology Management is a standalone product that requires no other SAS modules.

Client environment

- Linux for x86 (x86-32): RHEL 4, SuSE SLES 9
- Microsoft Windows (x86-32 and x64): Windows XP Professional, Windows Server 2003 family, Windows Vista*

Server environment

- AIX: Versions 5.3 and 6.1 (x64) on POWER architectures
- HP-UX Itanium: HP-UX 11iv2 (11.23), 11iv3 (11.31)
- HP-UX PA-RISC: HP-UX 11iv2 (11.23), 11iv3 (11.31)
- Linux for x86 (x86-32): RHEL 4, SuSE SLES 9
- Linux for x64 (EM64T/AMD64): RHEL 4, SuSE SLES 9
- Microsoft Windows (x86-32) and (x64): Windows XP Professional, Windows Server 2003 family, Windows Vista*
- Microsoft Windows on x64 (EM64T/AMD64): Windows XP Professional for x64, Windows Server 2003 for x64, Windows Vista* for x64
- Solaris on SPARC: Versions 9 and 10
- Solaris on x64: Version 10

* NOTE: Windows Vista Editions that are supported include Enterprise, Business and Ultimate

Required software

- SQL database (i.e., MySQL or Microsoft SQL Server)

Key Features (continued)

Builds subject-matter expertise into search-and-retrieval activities

- Lets you create an ontology that establishes and organizes the hierarchical relationships of semantic terms using advanced linguistic technology.
- Identifies the explicit and implicit connections (i.e., structural relationships) between document attributes.
- Identifies classes and subclasses of information within hierarchies.
- Lets you define attributes associated with each class and subclass within the hierarchy.
- Automatically applies instantiation definitions to appropriate content when information is sought through search or retrieval, regardless of document location or language.

Consistently applies subject-matter expertise in real time

- A predetermined list of semantic terms helps organizations begin their ontology development.
- Easy-to-use GUI enables users to add semantic terms to customize relationships among terms.
- Relevance and instances of occurrence can be specified to consistently identify relationships among areas of interest. The resulting hierarchies can be used as input taxonomies to automatically categorize content using SAS Enterprise Content Categorization.
- Content can be associated per the defined ontology at run time, creating metadata when the ontology is connected to a database using the product APIs.
- Advanced browsing of data is supported so that key document attributes can be clearly understood.
- A query window supports both simple and complex (linked simple queries) to find desired classes, drill into their details and replace associated information as needs change.
- Hierarchy and inverted hierarchy views facilitate parent-child relationship understanding and can be used as taxonomies for downstream applications.
- The ontology management server provides the ability to show users the results by applying the ontology to document collections in real time.

Centralizes administration for collaborative ontology development and maintenance

- Ontologies are retained as projects that can be saved, shared and collaborated on for ongoing maintenance and updates, with users' rights controlled by administrative authorizations.
- Default permissions can be set at each of the three project levels.
- Administrators can specify individuals as members of groups with collectively defined permissions.
- Users can be limited to read-only access, or they can have the ability to add, delete or modify classes, attributes (aka "slots") or instances.
- Supports class- and instance-level version control.
- Provides find-and-replace capabilities to retain consistency with broad changes.
- Provides report logging that records what instance-level changes were made by whom as part of the standard audit functions.
- Semantic terms are systematically managed and the conditions and connections associated with isolated types of key information are defined so those defined as most relevant are delivered to the right person.



SAS Institute Inc. World Headquarters +1 919 677 8000

To contact your local SAS office, please visit: www.sas.com/offices

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2010, SAS Institute Inc. All rights reserved. 104380_553142.0210