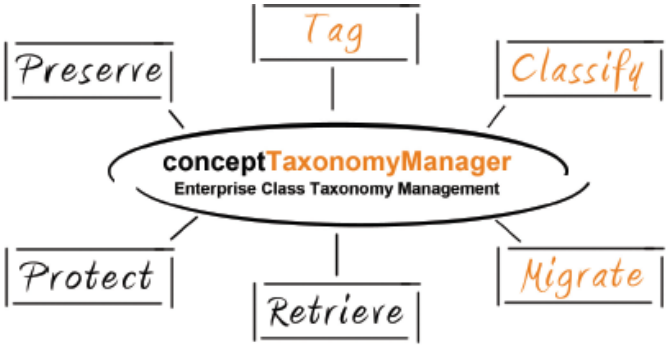
conceptSearching

conceptTaxonomyManager Product Overview



conceptTaxonomyManager is a major component in Concept Searching's **Smart Content Framework™** for information governance, which was developed as a toolset that provides the enterprise framework to mitigate risk, automate processes, manage information, protect privacy, and address compliance issues. The framework is being used to improve search, in records management, enterprise metadata management, compliance, migration, privacy and governance.

The Technologies

The technologies incorporate automated classification and taxonomy tools to manage unstructured enterprise content. Utilizing Concept Searching's unique compound term processing, the technologies deliver a set of outcomes that are not achieved by any other classification engine.

Compound term processing means that Concept Searching's statistical engine can understand, out of the box, the incremental value of keywords, multi-word fragments and compound terms, and as a result identify concepts resident within an organization's own information repositories that are highly correlated to particular topics.

With the identification of these highly correlated topics, in the form of keywords, multi-word fragments and compound terms, the result is automatically generated semantic metadata that is unique to that particular organization.

Utilizing compound term processing, the semantic metadata generated from the organization's own content can be used as clues to help automate building out the taxonomy. Content will be classified to one or more nodes in the taxonomy, or multiple taxonomies, based on the concepts within the content.

Product Description

conceptTaxonomyManager is a simple to use, intuitive user interface designed for Subject Matter Experts without IT or Information Scientist expertise to build, maintain and validate taxonomies for the enterprise. Features delivered include:

- Compound term processing technology that identifies 'concepts in context'
- Automatic semantic metadata generation as content is created or ingested
- Automated classification of content to one or more nodes in one or more taxonomies
- Aggregated content from multiple content sources
- Taxonomy management rapidly deployed and easily managed
- Controlled vocabularies
- Multiple taxonomy support
- Automatic taxonomy node clue suggestion
- Ontological features supported
- Dynamic screen updating to immediately see impact of changes in the taxonomy
- Document movement feedback to see cause and effect of changes without re-indexing
- Full security model enabling lock down of nodes, branches and complete taxonomies to particular users and or groups of users
- Supports rollback to previous state of the taxonomy
- Data held in standard SQL enabling BI tools to be layered over the data to build reports and dashboards
- Can be used by any search engine to improve findability during search
- SOA compliant, bowser based application
- Distributed taxonomy management feature to support locking/unlocking of nodes by taxonomy managers to work in a distributed manner

Features

Taxonomies and their classes are managed via the taxonomy menu and the management features are dependent on the permissions assigned to the user. Taxonomies are easily added via the taxonomy drop down list. Once the taxonomy is given a name an empty taxonomy will be created and selected, ready for editing.

The taxonomy is presented as a tree view, allowing for nodes, or classes, to be expanded out to browse and select child nodes. If multiple taxonomies are available, a taxonomy dropdown list will be displayed. Apart from the addition of taxonomies, all tree management is done via a context menu, which is displayed when right-clicking on a taxonomy or class node with the mouse as shown below. Drag and drop capabilities for moving or copying a node are also provided.

Industry standard formats and taxonomies such as OWL and MeSH can be easily imported as well as any organizationally defined taxonomy. Clues are used to describe the language found in documents that make them about a particular topic. The scoring of the clue can be manually entered or changed by the user. Negative clues can also be added to lower the ranking of documents that you do not want in the node.

Clue Types
The user can add terms manually, provide a score, make the clue mandatory, and select the Type. Clue Types supported:

- Standard – Single words and phrases
- Case Sensitive
- Document Metadata – Partial matching on metadata values
- Phonetic Matching – Ability to locate topics regardless of spelling variations
- Regular Expression – Any regular expression can be used, such as part numbers or credit card numbers
- Class ID – Can only be classified if Parent or Grand Parent is classified
- Language Filters – Individual topics limit individual clues to specific languages

Suggest Clues for Class
Although manual entry of clues for a class is available, typically the end user may not think of all clues. A 'Suggest Clues for Class' feature is available that will search the document corpus and identify documents that are about the new node to provide the end user with additional choices to include or exclude clues from the node.

conceptTaxonomyManager will search the document corpus for similar terms and concepts and then displays these as clues. Utilizing compound term processing, clues returned can be multi-word terms that identify a concept, as well as single words and acronyms.

Clue Feedback
In some instances a user may want to understand why a clue was automatically suggested. By clicking on the clue a summary and extract is displayed that shows the clue in context of the document. The end user can keep the clue, remove it, or modify the weighting.

Document Movement Feedback
Powerful functionality is available to tune the taxonomy classifications. Document Movement Feedback provides the mechanism to evaluate the changes on the taxonomy in real time without the need to reclassify the content. This on the fly feature will display the new classifications based on changes made to the scores. Indicators show how the score changes will impact the classification.

Indicators include: Document remains classified with a higher score; Document remains classified but with a lower score; Document remains unclassified and the score does not change; Document will now become classified; Document either stays or becomes unclassified.

Auditing Features
The product will track all changes and provides a rollback feature to undo unapproved changes.

Distributed Taxonomy Management
This feature is a requirement for organizations that have many taxonomy operators, extremely large collections of documents, and where taxonomy management is a critical business process. This feature can be implemented on any number of servers, and several taxonomy managers can be assigned to a server to ensure the level of throughput needed. Real time locking mechanisms are used to make nodes of the taxonomy inaccessible to other taxonomy managers while the node is being edited. The taxonomy managers can see when a node is locked and who has locked it, as well as when it becomes available.

The Distributed Taxonomy Management feature is totally transparent to the end user and all locking and unlocking of the nodes by the taxonomy managers are coordinated by the central server. The portion of the taxonomy that is being updated by a taxonomy manager will appear in red. It will appear in gray to other users to indicate that it is currently locked.

Requisite Products: None

Optional Components: conceptTaxonomyWorkflow

Americas
+1 703 531 8567
info-usa@conceptsearching.com

Europe
+44 (0)1438 213545
info-uk@conceptsearching.com

Canada
+1 703 531 8567
info-canada@conceptsearching.com

Australia
+61 (0)2 8006 2611
info-australia@conceptsearching.com

New Zealand
+64 (0)4 889 2867
info-nz@conceptsearching.com

Africa
+27 (0)21 712 5179
info-sa@conceptsearching.com

Marketing and PR
International: +1 703 531 8564
Europe: +44 (0)1438 213545
marketing@conceptsearching.com

Follow us on Twitter
@conceptsearch

www.conceptsearching.com