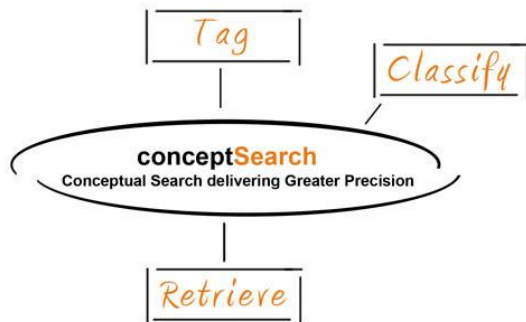


conceptSearching

conceptSearch Product Overview



conceptSearch is incorporated into 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.

Product Description

conceptSearch is a unique, language independent technology. It is the first content retrieval solution to integrate relevance ranking based on the Bayesian Inference Probabilistic Model, concept identification based on Shannon's Information Theory, and probabilistic latent semantic indexing.

Unlike other enterprise search engines, that require significant customization with marginal results, conceptSearch is delivered as an out of the box application that demonstrates a simple search interface and indexing facilities for internal content, web sites, file systems and XML documents. Application developers experience a minimal learning curve and the organization can look forward to a rapid return on investment.

Because of the innovative technology, conceptSearch delivers both high precision and high recall. Precision and recall are the two key performance measures for information retrieval. Precision is the retrieval of only those items that are relevant to the query. Recall is the retrieval of all items that are relevant to the query. Yet most information retrieval technologies are less than 22% accurate for both precision and recall. The ideal goal is to have these facilities balanced. *Compound term processing* has the ability to increase precision with no loss of recall.

This is particularly important for organizations that need sophisticated search and retrieval solutions. By weighting multi-word phrases, instead of single words, or words in proximity, the retrieval experience is more accurate and relevant. The ability for the search engine to identify concepts enables organizations to improve the search experience for a variety of business requirements.

Compound Term Processing

Unique to Concept Searching technology, *compound term processing* identifies and forms compound, or multi-word, terms, typically consisting of two to five words. Search can be performed with a greater degree of accuracy because the ambiguity in single words is no longer a hindrance.

For example, 'triple' is a single word term but 'triple heart bypass' is a compound term that provides a more granular meaning. A search for 'survival rate after triple heart bypass' will locate documents about this topic even if the precise phrase is not contained in any of the documents.

Required Products

conceptClassifier has the automatic classification process that, during indexing, identifies categories that each document belongs to. Each category is identified by a unique descriptor and is associated with key descriptive words and/or phrases held in the database. This approach enables a rapid implementation of a corporate taxonomy with all documents classified to multiple nodes at index time. Ideally, the taxonomy can be used to browse the document collection, or as a filter when running ad hoc searches. Classification of internal and external repositories can be done in real time or scheduled as required.

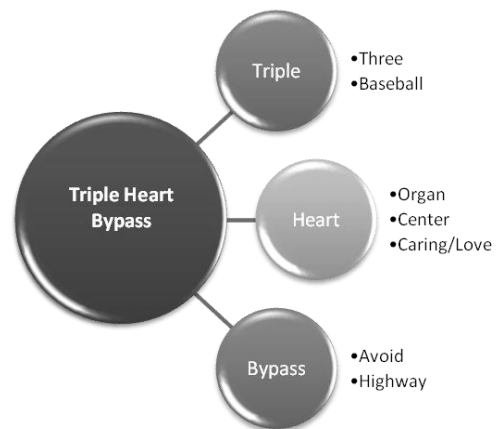
conceptTaxonomyManager is an easy to use, yet powerful, taxonomy tool that creates the framework to classify content based on concepts to one or more nodes in the taxonomy. Features to enable Subject Matter Experts to interact with the taxonomy can simplify ongoing maintenance. For example, automatically generating compound term clues from the document corpus, dynamically showing the effect of changes on the taxonomy, and class weighting influenced by parent, child and sibling can reduce taxonomy development and ongoing maintenance by 66%-80%.

conceptSearching

Product Features

The technology can isolate the key meaning that is normally expressed as proper nouns, nouns phrases and verb phrases. Although linguistic products can do this, their performance is highly variable depending upon the vocabulary and language in use. Concept Searching technologies are based on a statistical language independent model that can accept queries in natural language with the user typing words, phrases or whole sentences. The system then analyzes the natural language query to extract the keywords and phrases to identify the main concepts and retrieve content that is highly relevant.

- Compound terms are extracted when content is indexed, enabling the delivery of relevant content at the top of the search results.
- Relevance ranking display extracts from the documents based on the query are returned to the user.
- Search refinement delivers to the end user highly correlated suggested concepts that may be used to refine the search. Taxonomy browse capabilities are also standard.
- Documents can be classified into one or more taxonomy nodes, enhancing the precision of documents returned.
- In addition to static summaries, Dynamic Summarization, a modified weighting system, can be applied that will identify real time short extracts that are most relevant to the user's query.
- Taxonomy and faceted navigation.
- Text preview capability of attachments such as email or pdf files without having to open the originating application. Search results will be highlighted in the attachments.
- Related topics will return results based on the conceptual meaning of the search terms used.
- Based on previous queries, or on extracts retrieved, end users can use the text to perform additional searches to retrieve more granular results.
- Presents a single integrated view of content regardless of where it resides.



The search results using *compound term processing* will return documents, even if the exact terms are not contained in the document i.e. coronary artery surgery, heart surgery.

Standard Features

- Related Topics
- Query Expansion
- Filtering and Clustering
- Language Stemming
- Diacritic Support
- Stoplists
- Stemming, Fuzzy Stemming
- Soundex

Technology

The product is based on an open architecture with all APIs based on XML and Web Services. Transparent access to system internals including the statistical profile of terms is standard.



Microsoft Partner

Gold Independent Software Vendor (ISV)
Silver Content Management
Silver Search
Silver Portals and Collaboration

© 2012 Concept Searching

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

