

Taxonomy in Text Archives & Beyond  
*Special Libraries Association, News Division*  
June 14, 2006



LexisNexis® SmartIndexing Technology™ is a proprietary rule-based automated classification system that adds structure to the billions of documents available through the LexisNexis services. Applied against news, business and legal content, SmartIndexing taxonomies include tens of thousands of index terms for companies, industries, subjects, geographic locations and legal topics. Built and maintained by information professionals, the rule sets, thesaurus metadata and taxonomies ensure that customers return relevant, actionable results with every search. News and business index terms are available in US/UK English, German and French. New terms are added monthly and the entire archive of news content is reindexed at least once a year.

LexisNexis product interfaces – including the new *Nexis*® service – rely on SmartIndexing to offer tools and enhancements for both end users and power users. From simple pick lists to navigable hierarchies to answer set analysis, SmartIndexing enables the features that save users time and help them to discover the right answers – fast.

LexisNexis also licenses SmartIndexing Technology to customers for application against their internal content repositories. Customers can select from any of the pre-existing taxonomies or develop custom terminology.

For further information, contact your LexisNexis account representative or Teresa MacGregor, Taxonomy Product Consultant ([teresa.macgregor@lexisnexis.com](mailto:teresa.macgregor@lexisnexis.com)).



Access Innovations, Inc., focused on efficient and precise information storage and retrieval since 1978, has applied its Data Harmony software to create NewsIndexer, a tool invaluable to a news organization. NewsIndexer helps reporters, editors, and archivists assign relevant, consistent indexing terms to news articles, and helps searchers find those materials with precision.

Written in Java, NewsIndexer combines:

- a specialized and extensive taxonomy representing news concepts
- a rulebase governing the intelligent application of indexing terms
- authority files with names and geographic locations, optional

Integrated into the workflow, it suggests indexing terms for an article or blog entry or assigns descriptors to archived items, by single document or in batches.

Indexing terms are drawn from a controlled vocabulary representing the concepts in your news

articles. The rulebase provides instructions for when a word has multiple meanings, or when a concept is expressed in varying forms. NewsIndexer's rules are automatically generated and easily finetuned by your people or at Access Innovations. Both the terms and rulebase can be modified to keep up with changing usage or evolving language.

Flexible and customizable, NewsIndexer provides relevant and complete indexing for each article, leading to more productive retrieval.

NewsIndexer is available through Access Innovations, Inc., [www.accessinn.com](http://www.accessinn.com), (505) 998-0800.



Factiva Intelligent Indexing™ is a set of company, subject, industry and region taxonomies. It is implemented universally across the entire Factiva content set, allowing users to find more relevant information, quickly.

This comprehensive and award-winning taxonomy has been developed over the last 25 years. It is a proprietary system, building on the strengths of its parent company schemes - Dow Jones and Reuters, and routinely incorporating customer feedback. The taxonomy is poly hierarchical and is applied to content in all 22 languages. Factiva Intelligent Indexing™ codes are reviewed once a quarter; all new terms and changes are published in advance of implementation.

The Factiva Intelligent Indexing taxonomy can also be licensed for incorporating into a company's own taxonomy, for indexing information other than Factiva content and for integrating into a portal or intranet. Factiva's Taxonomy Consultants help organizations to build, deploy and maintain working taxonomies tailored to their individual needs, often in conjunction with Factiva Synaptica vocabulary and metadata management software.

<http://taxonomywarehouse.com> is known as the most complete independent, trusted, directory and source of over 530 taxonomies, thesauri, classification schemes and other authoritative files about taxonomy-related standards, resources and events.

For further information, contact your Factiva representative or [beth.golden@factiva.com](mailto:beth.golden@factiva.com)