



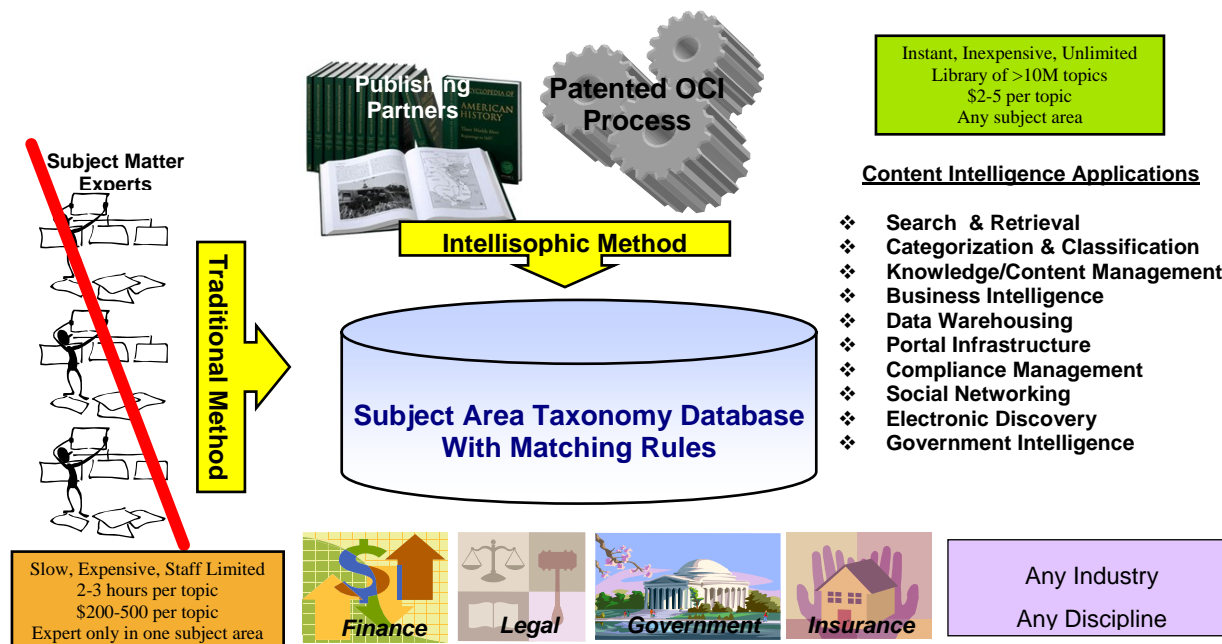
## **A Guide to Concept Taxonomies**

**An Intellisophic Proprietary White Paper**

**Taxonomic content is the key to next generation Content Intelligent solutions.** Intellisophic provides taxonomic content to fuel content intelligence applications. Taxonomic content is critical to enterprise applications which leverage knowledge trapped in unstructured data such as documents, emails, corporate correspondence, reports, memos, call center notes, CRM notes, contracts, IRC chatter, instant messages, web pages and other text files. This industry sector includes applications such as search and retrieval, content categorization, business and government intelligence, data warehousing, email archiving, electronic discovery, social networking, compliance, document forensics and other information management tools. Experts proclaim the \$8-10B content intelligence market is one of the fastest growing segments of enterprise technology.

Intellisophic has developed a disruptive technology for creating deep, authoritative reference content on any subject area. Unlike other methods for taxonomy development which are limited by the expense of corporate librarians and subject matter experts, Intellisophic content is machine developed, leveraging knowledge from the world's most respected reference works. Through this patented process, the largest library of taxonomic content has been created, covering several million topics defined by hundreds of millions of terms. This taxonomy library is constantly growing with the addition of new titles and publishing partners. Currently contributing to the library are esteemed reference publishers like John Wiley & Sons, Thompson Gale, Pharmaceutical Press, Woodhead Publishing, WorldBook Encyclopedia, Oxford University Press, C. Hurst and Company, Aspen Publishers and American Society of Microbiology.

## TAXONOMY METHODOLOGIES



The taxonomies are delivered in XTM Topic map format, which is an open XML based format; RDF, which is a relational database; or OWL Web Ontology Language, which was designed specifically for applications that need to process content. Intellisophic taxonomies are also customized to fit within specific vendor applications. These taxonomies provide a table of contents for navigating the concept tree as well as concept signatures (also referred to as matching rules) that are essential for identifying all concept instances in a text document. In other words, the taxonomies are rich with content, organized for effective usage and easily integrate into existing applications.

**Deep, rich taxonomies with matching rules make Intellisophic taxonomies unparalleled.** The abundance of subject areas covered can be overwhelming. For this reason, we have developed a process to work with our customers which ensures the appropriate taxonomies are chosen. First, the customer speaks with Intellisophic to identify their areas of interest. Subject areas can be very broad so this process allows customers to clearly define their interests. Once this is done, Intellisophic will search through its vast library and pull all relevant taxonomies. This results in a preliminary recommendation, or a high level overview, that will often contain an abundance of taxonomies. The customers are provided with descriptions and topic counts so they can easily pick the taxonomies that are most appealing. Intellisophic will review the customer selection to make certain that there are no gaps in coverage. Intellisophic will then pull the taxonomies from their library on hand or ingest the necessary content. The resulting taxonomies will be formatted to work within the customer's application. The taxonomies are then deployed to the customer and fine-tuned if necessary.

The following is a high level catalog of Intellisophic taxonomies. These are subject area categories and sub-categories. Each level will undoubtedly contain multiple taxonomies. Intellisophic's library is constantly growing so if an area of interest is not listed below, please contact us to inquire about it.

### **Arts & Humanities**

- **Architecture & Design**
- **Fine & Performing Arts**
  - **Painting & Sculpture**
  - **Music & Dance**
  - **Theatre & Film**
- **Language & Linguistics**
- **Literature**

### **Business**

- **General**
- **Accounting**
  - **Accounting**
  - **Auditing**
  - **Taxation**
  - **Credit & Collection Law**
- **Advertising & Marketing**
- **Economics**
- **Finance**
  - **General**
  - **Banking & Mortgage**
  - **Estate Planning**
  - **Insurance**
  - **Investment**
  - **Regulation**
- **Management**
- **Non-profit Organizations**

### **Energy**

- **General**
- **Coal**
- **Electricity**
- **Oil & Gas**
- **Nuclear**
- **Water**

### **Engineering & Materials**

- **General Engineering & Materials**
- **Mechanical Engineering**
  - **General**
  - **Welding & Fatigue**
- **Civil & Structural Engineering**
  - **General**
  - **Construction**
- **Industrial Engineering**
- **Environmental Engineering**
- **Electrical Engineering**
  - **General**
  - **Communications**
  - **Electronics**
- **Metallurgy**
- **Nuclear Engineering**
- **Plastics & Polymers**

### **Environment**

- **General**
- **Agriculture**
- **Air**
- **Geology**
- **Ecology**
- **Food Science**
- **Forestry, Fishing & Hunting**
- **Geography**
- **Water & Soil**
- **Weather**
- **Animals & Wildlife**

### **Hospitality**

- **General**
- **Food & Beverage**
- **Lodging**
- **Travel & Tourism**



### Information Sciences

- *Broadcasting*
- *Computer Sciences*
- *Internet*
- *Motion Picture & Sound Recording*
- *Publishing*
- *Telecommunications*

### Law, Government, Crime & Justice

- *General Law & Criminal Justice*
- *Insurance Law*
- *Medical Law*
- *International Law*
- *Criminal Law*
- *Energy Law*
- *Environmental Law*
- *Labor Law*
- *Civil Law*
- *Commercial and Economic Law*
- *Constitutional Law*
- *Cyber Law*
- *Property and Trust Law*
- *Social Security Law*
- *Sports Law*
- *Telecommunications Law*
- *Transport Law*
- *Corporate Law*
- *Financial Law*
- *Government*
  - *Intergovernmental Organizations*

### Life Sciences, Medicine & Healthcare

- *General*
- *Addictions*
- *Alternative Medicine*
- *Associations*
- *Bacteriology*
- *Biology*
- *Biotechnology*
- *Critical Care*
- *Dentistry*
- *Dermatology*
- *Diseases*
- *Drugs & Pharmacy*
- *Ethics*
- *Genetics*
- *Manufacturing*
- *Microbiology*
- *Neurology*
- *Nursing*
- *Nutrition*

### Life Sciences, Medicine & Healthcare

(continued)

- *Physiology*
- *Psychiatry*
- *Radiology*
- *Regulations*
- *Safety*
- *Surgery*
- *Veterinary Medicine*
- *Virology*

### Manufacturing

- *Automotive*
- *Electronics*
- *Food, Beverage & Tobacco*
- *Textiles & Leather*
- *Wood & Paper Products*
- *Chemical, Plastics & Rubber*
- *Nonmetallic Metals*
- *Metals & Machinery*
- *Computer & Electronic Products*
- *Household*

### Mathematics & Science

- *General*
- *Astronomy*
- *Chemistry*
  - *General*
  - *Analytical Chemistry*
  - *Biochemistry*
  - *Inorganic Chemistry*
  - *Industrial Chemistry*
  - *Electrochemistry*
  - *Environmental Chemistry*
  - *Organic Chemistry*
  - *Physical Chemistry*
  - *Spectroscopy*
- *Earth Science*
- *Mathematics & Statistics*
- *Materials Science*
- *Nanoscience & Nanotechnology*
- *Physics*
  - *General Physics*
  - *Atomic, Molecular, & Optical Physics*
  - *Lasers & Optics Physics*
  - *Nuclear & High Energy Physics*
  - *Solid State Physics*
  - *Thermal Physics & Statistical Mechanics*
  - *Quantum Physics*
- *Sensor Materials*



### **Social Sciences**

- *General*
- *Anthropology & Archaeology*
- *Education*
- *Income & Demographics*
- *Nation & World*
- *Religion*
- *Philosophy*
- *Political Science*
- *Psychology*
- *Sociology*

### **Terrorism & War**

- *General*
- *Medical Treatment*
- *History*
- *Weapons & Warfare*
  - *Biological & Chemical Warfare*
  - *Guns*
  - *Knives*
- *Critical Infrastructures*
- *Organizations*
- *Financing*

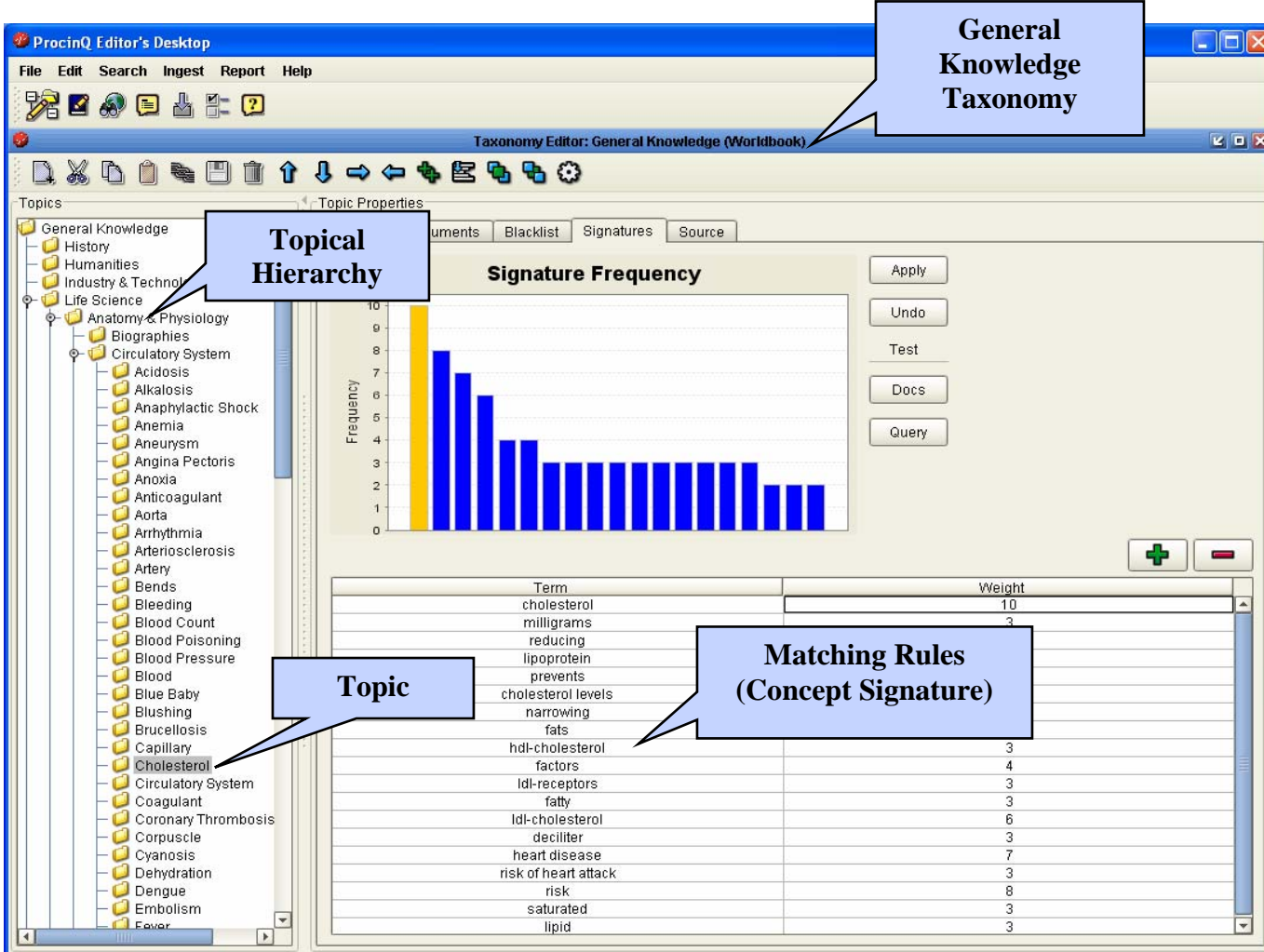
### **Transportation**

- *General*
- *Air & Aerospace*
  - *Air & Spacecraft*
  - *Radar*
  - *Engineering*
- *Land*
  - *Automotive*
  - *Trucking*
  - *Rail*
- *Sea*

### **General Knowledge**

- *Abbreviations & Acronyms*
- *Associations*
- *Directories*
- *Thesauri*
- *Miscellaneous*

Below is a screen view of a general knowledge (reference) taxonomy opened into the life sciences subject area, featuring cholesterol as a topic with associated matching rules.



**General Knowledge Taxonomy**

**Topical Hierarchy**

**Topic**

**Matching Rules (Concept Signature)**

Term	Weight
cholesterol	10
milligrams	3
reducing	
lipoprotein	
prevents	
cholesterol levels	
narrowing	
fats	
hdl-cholesterol	3
factors	4
ldl-receptors	3
fatty	3
Coronary Thrombosis	6
deciliter	3
heart disease	7
risk of heart attack	3
risk	8
saturated	3
lipid	3

By means of patents and reference publisher agreements, Intellisophic is the only company that is able to create taxonomic content through this method. With a growing library of content, the company is already offering a library that is orders of magnitude larger than any competitors in this space. It is the size and depth of this library that will make Intellisophic the de-facto solution for taxonomic content in the content intelligence space.