BioCAID™
Bioinformatic Conceptualization for
Automated Intelligent Discovery

Text mining is the science of processing and analyzing text collections to discover potentially useful and interesting knowledge. Discovering hidden and new knowledge relationships from scientific literature using the text mining methods is termed as literature-based discovery.

The rapidly growing text collections, such as the MEDLINE® database of over 28 million documents, and well defined ontologies provided by the Medical Subject Headings® and Unified Medical Language System® (UMLS) make biomedical literature an active and attractive domain for literature-based discovery.

Biomedical research is divided into highly specialized fields and subfields, with poor communication between them. The rate of growth of publications makes it difficult for a researcher to derive connections between Concepts from different research specialties. BioCAID's IMDI Profiler excels at quickly mining hidden connections among biomedical Concepts from large amounts of scientific literature and presenting them in as an intuitive visualization and navigation application.

Exploratory Knowledge Discovery

In order to perform exploratory discovery and generate hypotheses with respect to a Concept (say, A), we begin by creating an association network for Concept A. The association network consists of all the Concepts that co-occur with A, which are called the 'B Concepts'. The network is then expanded to include all of the co-occurring Concepts, C, for each of the Concepts in the set B. Each of those (C Concepts) is indirectly connected via at least a specified minimum number of intervening B Concepts.
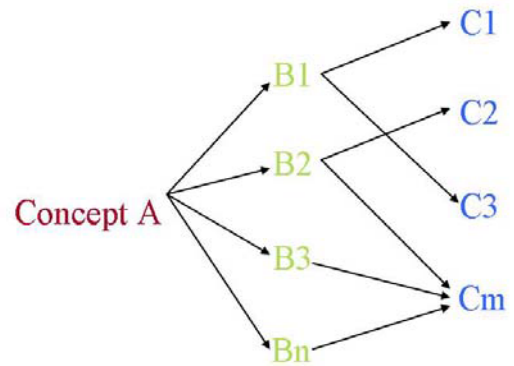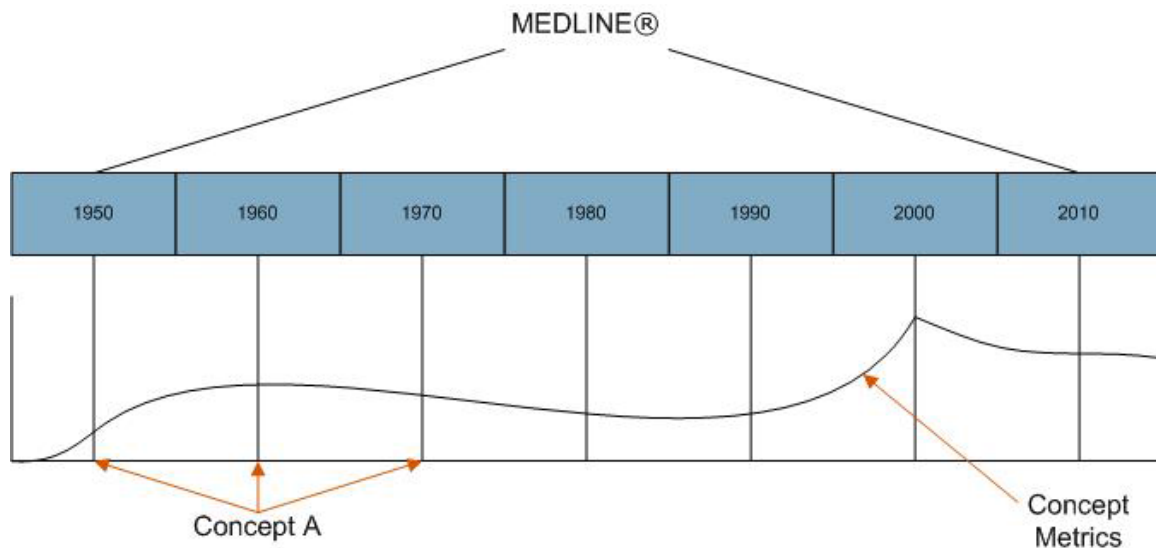

Confirmatory Hypotheses Research

Confirmatory Hypotheses Research is the process of determining the existence or strength of a relationship between two Concepts. First, the researcher would provide both an A and C Concept from a hypothesis proposing a relationship between the A and C Concept. Then, innovative relationship discovery algorithms are applied to the literature set to confirm whether a potential connection exists between A and C, due to the presence of a sufficient number of intervening B Concepts that strongly co-occur with the two input Concepts, A and C.
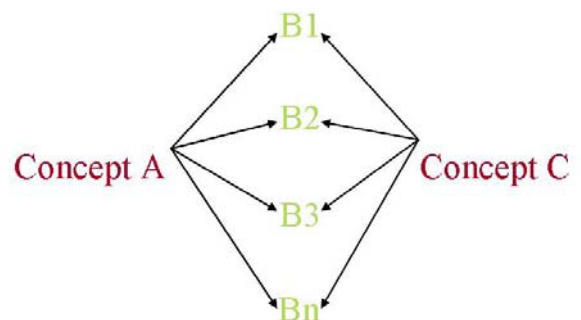
Predictive Knowledge Trending

Predictive Knowledge Trending is a method of profiling term and Concept metrics between equally divided time slices of the literature set. First, the published literature is divided into equal time periods. Then the same analysis (term occurrences, term relationships, and the types of relationships themselves) is applied to each time period. Finally, the trends in metrics are recorded and compared.

References for: **Literature-Based Discovery**

http://www.nlm.nih.gov/pubs/factsheets/medline.html

http://www.infovision.org.in/2007/topic/presentations/Prof.%20Vijay%20V.%20Raghavan.ppt

Find the Hidden Relationships in Biomedical Research

To walk through a demonstration video of an Alpha prototype from our prior research, go to www.biocaid.com and click DEMO. In this self-guided demo, you will be able to see how the functionality of the IMDI Profiler organizes and presents relationships between publications and terms to support the visualization, navigation and integration of Concepts and researcher interests in the MEDLINE Library.

If you have questions about system features or our company, please contact BioCAID via e-mail: info@biocaid.com