

Taxonomies and Ontologies for Graph Search

Life Science and Healthcare organizations use RDF/SKOS/OWL based vocabularies, thesauri, taxonomies and ontologies to organize enterprise knowledge. There are many ways to use these technologies but one that is gaining momentum is to semantically index unstructured documents through ontologies and taxonomies.

In this talk we will demonstrate two projects where we use a combination of SKOS/OWL based taxonomies and ontologies, entity extraction, fast text search, and Graph Search to create a semantic retrieval engine for unstructured documents.

The first project organized all science related artifacts in Malaysia through a taxonomy of scientific concepts. It indexed all papers, people, patents, organizations, research grants, etc, etc, and created a user friendly taxonomy browser to quickly find relevant information, such as, "How much research funding has been spent on a certain subject over the last 3 years and how many patents resulted from this research".

The second project discusses a large socio-economic content publisher that has millions of documents in at least eight different languages. Reusing documents for new publications was a painful process given that keyword search and LSI techniques were mostly inadequate to find the document fragments that were needed. Fortunately the organization had begun developing a large SKOS based taxonomy that linked common concepts to various preferential and alternative labels in many languages. We used this taxonomy to index millions of document fragments and we'll show how we can perform relevancy search and retrieval based on taxonomic concepts.

[View a recording of the event here](#)

Download the presentation slides [here](#)