Semantic retrieval

Provides a semantic search mechanism

AIM AND FUNCTIONALITIES

The purpose of the semantic retrieval module is to integrate an ontology-based Boolean query model with a vector-space Information Retrieval model, in such a way that queries can be entered in a formal ontology query language (such as SPARQL) and answered by returning the multimedia documents displaying information related to the formal query answer.

In addition to the multimedia content repository, the module assumes the existence of a domain knowledge base (KB) described by a domain ontology, and the availability of annotations of the content by KB entities. One of the fundamental novel features provided by the module is a ranking algorithm integrating the classic IR view of the notion of relevance and approximate search results, with the formal ontology-based notion of query answering.

PERFORMANCE

The retrieval module has been tested on a set of twenty example queries, compared to a conventional keyword-only search, using the Jakarta Lucene library. Although a systematic efficiency testing has not yet been conducted, the average informally observed response time on a standard professional desktop computer is below 30 sec. The measurements show a clear improvement by semantic retrieval with respect to a keyword-based engine.

INTEROPERABILITY

The retrieval module makes minimum assumptions about the content, ontology, and query formats. It admits any OWL or RDFS ontology, and puts no restriction on the content modality or format. The only assumptions are the availability of annotation links between the ontology and content segments.

OWNERSHIP

The owner of this component is UAM.