

2010 Semantic Technology Conference Proposal

Authors:

Bettina Schimanski and Jennifer Miller

Submission Date: January 18, 2010

Title:

Unhappy With Internal Corporate Search? Come Learn Tips And Tricks For Building a Controlled Vocabulary Ontology

Brief Description:

Unhappy with internal corporate search? One improvement is to implement a controlled vocabulary ontology. We'll present our work on the construction of an ontology from the content and technical perspectives.

Summary:

Are your employees unhappy with internal corporate search? Frequent complaints include: too many results to sift through; results are unrelated/updated; employees aren't sure which terms to search for. One way to improve intranet search is to implement a controlled vocabulary ontology. Employing this takes the guess work out of searching, makes search efficient and precise, educates employees about the lingo used within the corporation, and allows employees to contribute to the corpus of terms. It promotes internal corporate search to rival its superior sibling, internet search.

We will cover our experiences, lessons learned, and conclusions from implementing a controlled vocabulary ontology at Sandia National Laboratories. The work focuses on construction of this ontology from the content perspective and the technical perspective.

We'll discuss the following:

- The tool we used to build a polyhierarchical taxonomy
- Examples of two methods of indexing the content: traditional "back of the book" and folksonomy word-mapping
- Tips on how to build future search capabilities while building the basic controlled vocabulary
- How to implement the controlled vocabulary as an ontology that mimics Google's search suggestions
- Making the user experience more interactive and intuitive
- Sorting suggestions based on preferred, alternate and related terms using SPARQL queries

In summary, future improvements will be presented, including permitting end-users to add, edit and remove terms, and filtering on different subject domains.

Additional Information:

Our presentation will benefit multiple audiences as we will discuss the topic from a library science as well as technical perspective. In our own struggles, we wish we had had similar guidance and insight to learn from.

Presentation Type: Conference Session - Full 60 min.

Level of Audience: Technical – Introductory

Track: Enterprise Semantics

Treatment: How-To

Other Topics We'd Like To Hear About: What other library science professionals are doing, advanced Jena API How-Tos, and social networking within an enterprise

10 Keywords That Relate To Your Session: Controlled Vocabulary, Ontology, Polyhierarchical Taxonomy, Folksonomy, Search