

SMW+

The Enhanced Retrieval Extension

Philipp Zaltenbach

www.ontoprise.de

Feb 12, 2009

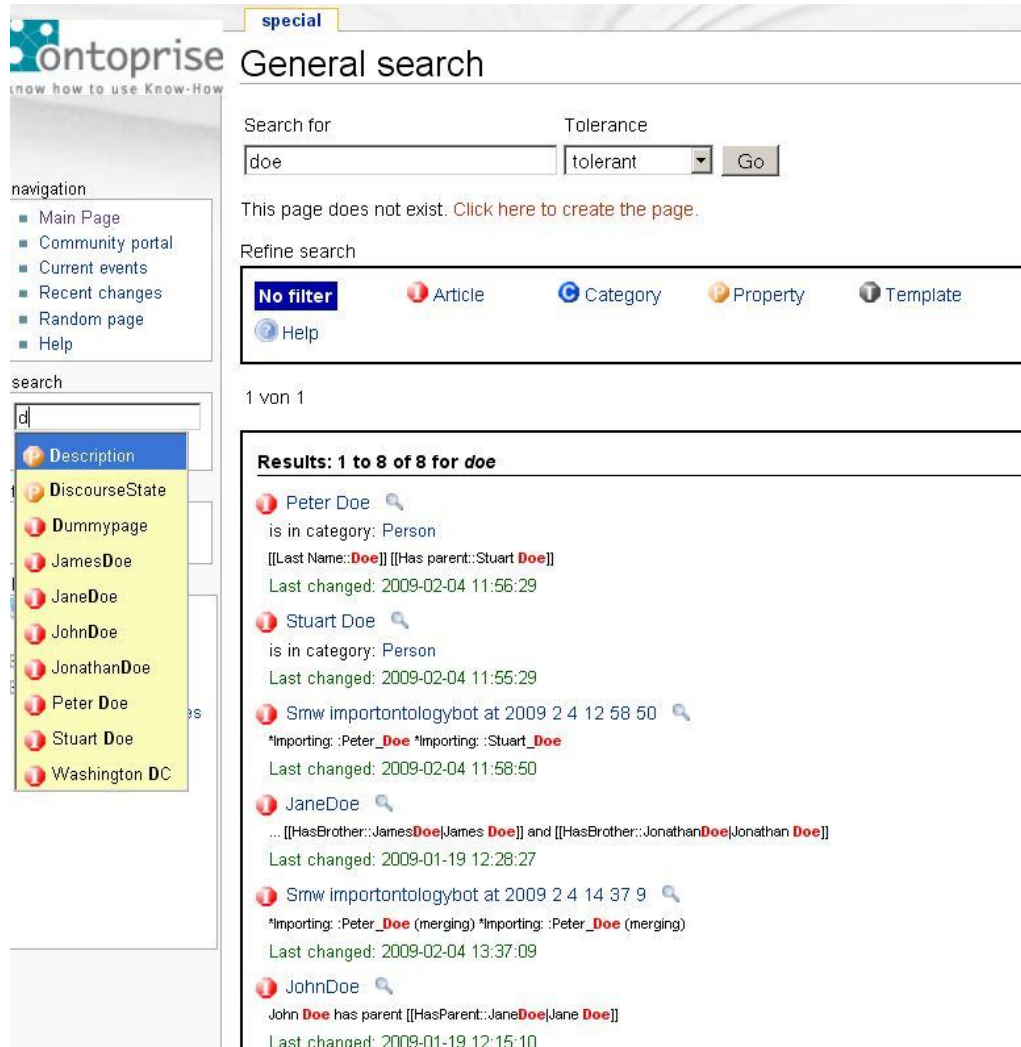
SemanticWiki mini-series session 5: "Semantic Wiki Applications & Use Cases (2):
horizontal applications"

Motivation for the Enhanced Retrieval Extension

- **Enhanced Retrieval Extension is based on the new MWSearch extension**

- **Advantages of MWSearch**
 - **“Did you mean?” (i.e. spell-checking and auto-suggest)**
 - **Fuzzy and wildcard queries**
 - **Wildcard search: *stan -> Kazakhstan, Tajikistan, Kyrgyzstan**
 - **Boolean Operators: AND, OR, NOT**
 - **Search for stop words & words with less than 3 characters (e.g. “VW”)**
 - **Prefix and Intitle search**
 - **Faster & Improved search results**
 - **Via better stemming, usage of related articles, sections, redirects**

Features of Enhanced Retrieval Extension



The screenshot shows the ontoprise search interface. At the top, there is a search bar with the text 'doe' and a 'Go' button. Below the search bar, there is a message: 'This page does not exist. Click here to create the page.' Underneath, there is a 'Refine search' section with buttons for 'No filter', 'Article', 'Category', 'Property', and 'Template'. The search results are displayed in a list format, showing 1 von 1 result. The results are for the query 'doe' and include:

- Peter Doe** (Person): is in category: Person. [[Last Name: **Doe**]] [[Has parent: **Stuart Doe**]]. Last changed: 2009-02-04 11:56:29.
- Stuart Doe** (Person): is in category: Person. Last changed: 2009-02-04 11:55:29.
- Srww importontologybot at 2009 2 4 12 58 50**: *Importing: :Peter_Doe *Importing: :Stuart_Doe. Last changed: 2009-02-04 11:58:50.
- JaneDoe**: ... [[HasBrother::JamesDoe|James Doe]] and [[HasBrother::JonathanDoe|Jonathan Doe]]. Last changed: 2009-01-19 12:28:27.
- Srww importontologybot at 2009 2 4 14 37 9**: *Importing: :Peter_Doe (merging) *Importing: :Peter_Doe (merging). Last changed: 2009-02-04 13:37:09.
- JohnDoe**: John Doe has parent [[HasParent::JaneDoe|Jane Doe]]. Last changed: 2009-01-19 12:15:10.

- **Auto-Completion**
- **Unification of full-text search and ontology search**
- **Leverage search by making use of semantic, i.e. "Query Expansion"**
- **Full-text search on uploaded DOC and PDF files**
- **Preview on search results**
- **Search statistics**

How does "Query Expansion" work?

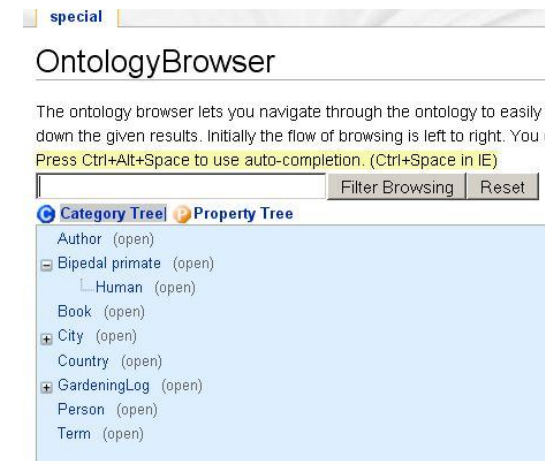
- Exploits the ontology in order to find related terms for entered search terms
- Related terms
 - Synonyms (Redirects, predefined properties: "also known as", "rarely labeled as")
 - Broader and narrower terms (Super- and sub-categories, predefined properties: "broader term", "narrower term")
- Search settings to control degree of Query Expansion
 - Settings: exact, semi-tolerant, tolerant
- Example
 - Entered search query *Ferrari F50* (Setting: tolerant)
 - Query gets expanded to
(*F50 OR Ferrari F50 OR Sportscar OR Gasoline Guzzler*)

How can a user insert term relations?

- Use the ontology browser
 - Model super- and subcategories, subproperties

- Semantic markup in wiki-text
 - E.g. On the article page of "Ferrari F50" use
 - [[Also known as::Sportscar]]
 - [[Rarely labeled as::Gasoline Guzzler]]

- Form-based input of related terms (for easy management of vocabularies / thesauri)



Edit Term: Porsche	
also known as:	<input type="text" value="Gas Guzzler"/>
rarely labeled as:	<input type="text" value="Steam hammer"/>
definition:	<input type="text"/>
example:	<input type="text"/>
broader term:	<input type="text" value="Sportscar"/>
narrower term:	<input type="text" value="Porsche GT"/>
Free text:	<input type="text"/>

Summary

- **The enhanced retrieval extension allows for a more flexible and tolerant search**
- **Supports collaborative vocabulary management**
- **Reuse stored vocabulary data by relying on SKOS W3C standard**
- **Extension will be released to the community in March '09**
- **Resources**
 - **MW Search Explanation**
(<http://en.wikipedia.org/wiki/Wikipedia:Searching>)
 - **SMW+ User forum**
(<http://smwforum.ontoprise.com>)