

DD2476 Search Engines and Information Retrieval Systems

Project 2: Better Query Completion

Contact: Simon Stenström, Findwise (simon.stenstrom@findwise.se, 073-616 35 34)

This project is worth 3 ECTS credits. This means that it is expected to require 80 hours of work for each person in the group. The project formulation, method, and results are presented in a report as well as in a poster session. For more details, look at the course homepage, under Project in the menu.

Problem

Query completion (also known as “type ahead” or “auto suggestion”) is a feature that provides the end user with help on with to search for given the start of a search term. The regular query completion is using the index to figure out smart words to suggest (for example, if I start writing “school a” and the index contain the phrase “school assignments” the query completion module will return this).

This may however not be enough. The more you know about the data source that you are searching within, the more complex queries you will write. If I know that I am looking for a document written by “Simon Stenström”, I may search for “school assignments author:(Simon Stenström)”. And for the more advanced part of this query, regular query completion will not help you.

Assignment

Your assignment would be to create a query completion module for Apache Solr¹ that helps the user with more complex searches as well as the easy ones. This can include suggesting field names (See example 1 below) and content within a field (See example 2 below). It could also easily be expanded into suggestion usual query syntax (See example 3 below).

Example 1) The user write “aut”. Expected result: “author:()” if there is a field in Solr called author.

Example 2) The user write “author:(Sim”. Expected result: “author:(Simon)” if “Simon” can be found anywhere in the field author.

Example 3) The user write “test ” (with a space as the last character). Expected result: “AND”, “OR”, etc.).

¹ <http://lucene.apache.org/solr/>

About Findwise

Findwise is a growing IT consultancy company, founded in 2005 by a team of experts from the enterprise search industry. The company currently employs about 90 people (January 2012) and have offices in Sweden, Denmark, Norway and Poland.

The project is meant to be fun but could possibly be used as a demo of what our customers could do with their data more than just making it searchable. There are some ideas on how to solve the problem in the text, but other solutions are also warmly welcome.

If you have any questions, don't hesitate to ask (in Swedish or English).