

Web searching exercise: Query formulation

Purpose

The purpose of this exercise is for you to implement the knowledge you gained at lectures on information searching. The emphasis of this exercise is on query formulation in Web search engines. Just as a comparison, we will start with a library catalogue (LIBRIS). Other aspect of searching (types of resources available, which one to choose, searching strategy etc.) you can practice at home, using exercises created to this purpose by Sara.

Search engines resources for the exercise

- 1) AltaVista. <http://www.altavista.com/>
- 2) Yahoo. <http://www.yahoo.com>
- 3) Scirus. <http://www.scirus.com/>
- 4) Clusty. <http://clusty.com/>

Assignments

- 1) Go to LIBRIS at <http://websok.libris.kb.se/websearch/form>. Choose **Boolean search**. Run the following queries and observe the number of and type of hits you get, and make sure you understand what it is that you are actually retrieving:
 - i. pollution
 - ii. pollut*
 - iii. air pollution
 - iv. pollution AND air
 - v. pollution OR air
 - vi. pollution NOT air
 - vii. "air pollution" (as a phrase)
 - viii. pollution air water
 - ix. pollution AND air AND water
 - x. pollution AND air OR water
 - xi. pollution AND (air OR water)
- 2) Open four Internet browsers, with one search engine in each. Try to find a section on help with how to perform searching in each database.

Note: It is always important to get acquainted with searching syntax and searching options in every search engine and database because each has different syntax and options.
- 3) Compare the search engines for the syntax they use, by looking at the most advanced searching interfaces:
 - a. types of Boolean operators supported
 - b. how the Boolean operators are represented (+, -, /, AND, OR, NOT...)
 - c. is phrase searching supported, and how to express it
 - d. is truncation supported
 - e. by which fields is it possible to limit searching (e.g. by domain, document format etc.)
- 4) Go to AltaVista. In the Advanced Web Search interface, search with Boolean expression, run the same queries as in 1), and observe the number of and type of hits you get, and make sure you understand what it is that you are actually retrieving.
- 5) Go to other search engines and try to find out how these queries would be formed in those search engines, and which interface can be used (simple, advanced, etc. or all of them).

For more information on successful searching in different search engines see
<http://searchenginewatch.com/facts/>.