DD2476 Search Engines and Information Retrieval Systems 9.0 credits

Sökmotorer och informationssökningssystem

This is a translation of the Swedish, legally binding, course syllabus.

If the course is discontinued, students may request to be examined during the following two academic years

Establishment

Course syllabus for DD2476 valid from Autumn 2016

Grading scale

A, B, C, D, E, FX, F

Education cycle

Second cycle

Main field of study

Computer Science and Engineering

Specific prerequisites

SF1604 Linear algebra, SF1901 Probability theory and statistics, DD1338 Algorithms and Data Structures, or corresponding courses.

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

After completing the course you will be able to:

* explain the concepts of indexing, vocabulary, normalization and dictionary in Information Retrieval,

* give an account of different text similarity measures, and select a similarity measure suitable for the problem at hand,

* define a boolean model and a vector space model, and explain the differences between them,

* implement a method for ranked retrieval of a very large number of documents with hyperlinks between them,

* evaluate information retrieval algorithms, and give an account of the difficulties of evaluation,

* give an account of the structure of a Web search engine.

Course contents

Basic and advanced techniques for information systems: information extraction; efficient text indexing; indexing of non-text data; Boolean and vector space retrieval models; evaluation and interface issues; structure of Web search engines.

Course literature

C. D. Manning, P. Raghavan and H. Schütze: Introduction to Information Retrieval, Cambridge University Press, 2008.

Examination

• LABA - Laboratory lessons, 6.0 credits, grading scale: A, B, C, D, E, FX, F

• LABB - Laboratory lessons, 3.0 credits, grading scale: A, B, C, D, E, FX, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

In this course all the regulations of the code of honor at the School of Computer science and Communication apply, see: http://www.kth.se/csc/student/hederskodex/1.17237?l=en_UK.

Ethical approach

- All members of a group are responsible for the group's work.

- In any assessment, every student shall honestly disclose any help received and sources used.

- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.