



DD2476 Sökmotorer och informationssökningssystem

9,0 hp

Search Engines and Information Retrieval Systems

Kursplan för DD2476 gäller från och med VT19

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

Utbildningsnivå: Avancerad nivå

Huvudområde: Datalogi och datateknik

Lärandemål

Efter genomförd kurs ska du kunna:

- * förklara koncepten indexering, vokabulär, normalisering och kodbok i informationssökning,
- * redogöra för olika avståndsmått för text, och välja ett avståndsmått som är lämpligt för ett givet problem,
- * definiera en boolesk modell och en vektorrumsmodell, och förklara skillnaden mellan dem,
- * implementera en metod för rankad sökning av ett mycket stort antal dokument med hyperlänkar,
- * utvärdera informationssökningsalgoritmer, och redogöra för svårigheter med utvärdering,
- * redogöra för en internetsökmotors uppbyggnad.

Kursens huvudsakliga innehåll

Grundläggande och avancerade tekniker för informationssökning: informationsextraktion; effektiv textindexering; indexering av icke-textdata; booleska modeller och vektorrumsmodeller för sökning; utvärderings- och gränssnittsfrågor; internetsökmotorers uppbyggnad.

Undervisningsspråk

Undervisningsspråk anges i kurstillfällesinformationen i kurs- och programkatalogen.

Behörighet

SF1604 Linjär Algebra, SF1901 Sannolikhetsteori och statistik, DD1338 Algoritmer och datastrukturer, eller motsvarande kurser.

Litteratur

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

Examination

- LABA - Laborationer, 6,0 hp, betygsskala: A, B, C, D, E, FX, F
- LABB - Laborationer, 3,0 hp, betygsskala: A, B, C, D, E, FX, F

I denna kurs tillämpas skolans hederskodex, se: <http://www.kth.se/csc/student/hederskodex>.