



LUNDS UNIVERSITET  
Lunds Tekniska Högskola

Kursplan för läsåret 2011/2012  
(Genererad 2011-08-31.)

---

## WEBB-INTELLIGENS OCH INFORMATIONSSÖKNING

EITN01

### Web Intelligence and Information Retrieval

**Antal högskolepoäng:** 7,5. **Betygsskala:** TH. **Nivå:** A (Avancerad nivå). **Huvudområde:** Teknik. **Undervisningsspråk:** Kursen kan komma att ges på engelska. **Överlappar följande kurs/kurser:** EIT031. **Valfri för:** C4, D4, D4ks. **Kursansvarig:** Universitetslektor Anders Ardo, Anders.Ardö@eit.lth.se, Inst för elektro- och informationsteknik. **Förkunskapskrav:** EDA011 eller EDA016 Programmeringsteknik. **Förutsatta förkunskaper:** FMA420 Linjär Algebra. **Prestationsbedömning:** Tentamen samt godkända laborationer och inlämningsuppgifter. **Hemsida:** <http://www.eit.lth.se/kurs/eitn01>.

#### Syfte

Kursens syfte är att öka förståelsen för metoder för informationssökning (Information Retrieval), strukturering, och kunskapsextraktion (text mining) speciellt från Internet-baserade källor.

#### Mål

##### *Kunskap och förståelse*

För godkänd kurs skall studenten

på egen hand kunna:

- identifiera och formulera problem inom området Webb-Intelligens
- klassificera problemens svårighetsgrad i förhållande till den egna kunskapsnivån
- förstå och använda metoder från området informationssökning
- analysera informationskällor för att skapa strukturerad information

##### *Färdighet och förmåga*

För godkänd kurs skall studenten

- kunna implementera intelligenta applikationer för kunskapsextraktion av medelhög komplexitet, med utnyttjande av metodologi från informationssökning

##### *Värderingsförmåga och förhållningssätt*

För godkänd kurs skall studenten

- visa prov på förmåga att hantera för studenten nya metoder och resultat

### **Innehåll**

*Informationssökning*: Basala begrepp för sökning och rankning, vektormodeller, tf-idf baserad relevans. System för informationssökning.

*Sökspråk*: Olika sökspråk för sökning i strukturerade databaser presenteras.

*Strukturerad information*: Indexering, sökning och relevansrankning av sökresultat. Exemplifieras med hjälp av sökningar i strukturerade databaser (SRU/CQL).

*"Feature extraction"*: Extrahera egenskaper för dokument.

Grundläggande metoder för kunskapsextraktion och klassificering (som neurala nätverk, Support Vector Machines m fl) gås igenom och experimenteras med. Utnyttja extraherade egenskaper för att implementera ämnes-klassifikation för textdokument.

*Prestanda*: Prestandamätetal som precision och "*recall*".

### **Litteratur**

Baeza-Yates, R, Ribeiro-Neto, B: Modern Information Retrieval.  
Addison-Wesley 1999. ISBN: 0-201-39829-X  
Artiklar och document från Webben.  
Kursnoter och laborationer.