



Kursplan för läsåret 2008/2009  
(Genererad 2008-07-17.)

---

## WEBB-INTELLIGENS OCH INFORMATIONSSÖKNING

EITN01

### Web Intelligence and Information Retrieval

**Antal högskolepoäng:** 7,5. **Betygskala:** TH. **Nivå:** A (Avancerad nivå).

**Undervisningsspråk:** Kursen kan komma att ges på engelska. **Överlappar följande**

**kurs/kurser:** EIT031. **Valfri för:** C4, D3. **Kursansvarig:** Anders Ardo,

Anders.Ardo@eit.lth.se, Inst för elektro- och informationsteknik. **Förkunskapskrav:**

EDA011 Programmeringsteknik eller EDA016 Programmering. **Förutsatta förkunskaper:**

FMA420 Linjär Algebra. EDA132 Tillämpad artificiell intelligens (läses med fördel

parallellt). **Prestationsbedömning:** 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
- förstå och använda grundläggande "Computational Intelligence" teknologi

##### *Färdighet och förmåga*

För godkänd kurs skall studenten

- implementera intelligenta applikationer för kunskapsextraktion av medelhög komplexitet, med utnyttjande av metodologi från "Computational Intelligence" och 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.

*"Computational Intelligence"*: Grundläggande metoder för kunskapsextraktion och klassificering (som neurala nätverk, Support Vector Machines m fl) går igenom och experimenteras med. Utnyttja extraerade egenskaper för att implementera ämnesklassifikation för textdokument.

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

### **Litteratur**

Ricardo Baeza-Yates, Berthier Ribeiro-Neto: "Modern Information Retrieval", Addison-Wesley 1999, ISBN: 0-201-39829-X.

Artiklar och document från Webben.

Kursnoter och laborationer.