

Om Findwise

Findwise är ett svenskt konsultföretag inriktat på tillämpad sökteknologi. Vi vet att Sök är ett utmärkt verktyg för att råda bot på informationsöverflöd hos stora organisationer, liksom att applikationer baserade på Sök kan förenkla vardagen för vanliga människor.

Vi har kort beskrivit några projekt som vi tror lämpar sig ypperligt för datastudenter vid KTH. Gemensamt för våra projekt är att Sök har en central roll. Från Findwise sida räknar vi inte med några förkunskaper när det kommer till sökimplementation utan hjälper gärna till för att ni ska komma igång på ett bra sätt.

Exemplen nedan ska ses som idéer och är inte tänkta att begränsa er. Har ni några andra idéer inom något av områdena nedan – Kontakta oss så pratar vi om det!

Mer information om Findwise finns på <http://www.findwise.se>.

Kontaktperson och handledare

- **Namn:** Simon Stenström
- **E-mail:** simon.stenstrom@findwise.se
- **Telefonnummer:** 073-616 35 34
- **Besöksadress:** Sveavägen 28-30 (T-bana Hötorget)

Vi vill gärna träffa er och diskutera projekten för att kunna ta tillvara på era idéer och besvara era frågor om projekten tidigt och har därefter gärna en löpande kontakt med er.

Date night-sök

Att kombinera sökmotorer med kartsystem för att bygga ett fullfjädrat geo-sök är en spännande utmaning. Genom Google Maps finns väldigt mycket information om vad man kan göra i en stad kopplat till en precis plats.

Er uppgift blir att skapa ett söksystem som riktar sig till att hitta det bästa området givet ett antal sökkriterier. Tanken är att detta ska kunna användas för att planera en helkväll på stan, där fokus ligger i att minimera hur långt man måste promenera för att ta sig mellan allt man vill göra under kvällen. Ett exempel på en sökning skulle kunna vara "libanesisk restaurang, bio, cocktailbar i Stockholm" och resultatet skulle kunna presenteras som en karta där det minsta området som tillgodoser alla önskemål är markerat. En utökning på systemet (om tid finns) skulle kunna vara att presentera flera olika alternativ.

Uppgiften kan brytas ner i ett antal delar som är relativt oberoende av varandra:

- Extrahera och indexera data
- Hitta minsta område, givet koordinatlistor
- Presentera resultatet

Mer information om utveckling mot Google Maps:

<http://code.google.com/intl/sv-SE/apis/maps/index.html>

Breaking news

Det genereras mycket ostrukturerad information idag i till exempel tidningar och sociala nätverk. Twitter är ett exempel på en sådan källa och är väldigt intressant eftersom källan uppdateras konstant och direkt, men ändå i många fall är taggad med så kallade hashtags. Även hashtags kan vara formulerade på ett fritt och ostrukturerat sätt, men användare taggar ofta sina tweets med hashtags som de själva sett tidigare.

En uppgift blir att bygga en applikation som crawlar Twitter för att hitta nya "viktiga" hashtags så fort de dyker upp. En viktig hashtag kan definieras som att den plötsligt ökar kraftigt i användning. En utökning på systemet (om tid finns) kan vara att identifiera nyckelord som är viktiga för händelsen genom att använda en entitetstaggare (t.ex. ANNIE, <http://www.aktors.org/technologies/annie/>).

Uppgiften kan brytas ner i ett antal delar som är relativt oberoende av varandra:

- Identifiera vanliga hashtaggar
- Identifiera förändringar i användandet
- Eventuellt Identifiera nyckelord i tweets givet en hashtag
- Presentera resultat

Mer information om utveckling mot twitter:

<http://dev.twitter.com/>

Presentation av offentlig data

Under de senaste åren har mer och mer myndighetsdata börjat göras tillgänglig för allmän konsumtion. 1 Juli 2010 infördes PSI, en lag som berör hur denna offentliga information får återanvändas. På <http://www.opengov.se/> finns mycket av denna data att hämta.

En uppgift blir att skapa en portal som på ett snyggt sätt presenterar olika grafer över vad myndigheter spenderar pengar på. En utökning på uppgiften (om tid finns) är att koppla informationen till aktuella nyhetsinslag (till exempel vara att koppla artikeln "Försäkringskassan kastar ut Logica" i samband med siffror över vad Försäkringskassan spenderar pengar på).

Uppgiften kan brytas ner i ett antal delar som är relativt oberoende av varandra:

- Hämta data
- Automatiskt generera diagram
- Eventuellt göra sökningar mot nyhetskällor
- Presentera resultatet på ett överskådligt sätt