

Pressemelding

Stor dag for framtidens Web!

LYSAKER 16. januar 2008 - I går kveld norsk tid fikk et nytt språk for webben stempel som industristandard [1]. Den nye standarden heter SPARQL og er et språk for å formulere spørsmål. Man vil bruke det til å spørre om å få data knyttet til bl.a. dokumenter på webben, abstrakte ting som en værtype og til fysiske enheter som for eksempel en person. Språket er allerede i bruk knyttet til medisinsk forskning, i oljeindustrien, og dessuten er Wikipedia nå søkbar.

Når vi gjør et søk på web, foregår det mest i form av tekstsøk. Noen har kanskje prøvd å bruke spesielle ord i sine søk; søker man på "bil og fly", forventer man å få artikler som inneholder noe om både biler og fly. På Google kan man sette et minus-tegn foran et ord hvis man vil at dokumenter med ordet ikke skal være med i resultatet.

Presise søk får man sjelden gjort på denne måten, og derfor har det blitt utviklet egne spørrespråk. I database-systemer bruker man slike språk, men de har gjerne vært begrenset til en database. For å søke på en strukturert måte på et system med web-skala har man derfor utviklet spørrespråket SPARQL.

SPARQL kan brukes til å søke i data som er formulert med språket RDF. Det som er tilgjengelig slik kaller vi "Semantic Web", eller den semantiske webben [2]. Eksempler på data som er formulert på denne måten er data fra Wikipedia og en stor database med stedsnavn, som man finner på henholdsvis DBpedia [3] og Geonames [4], men det er mye, mye mer, milliarder av datapunkter kan det søkes i. Man bygger også broer til andre teknologier, og med GRDDL blir strukturerte data fra vanlige websider også tilgjengelig. Dette er standarder som allerede er på plass.

Computas AS [5] er blant de norske firmaene som tar i bruk SPARQL fra dag en. På vegne av oljeindustrien har vi fulgt utviklingen lenge, og forskningen som har sin basis i SPARQL kan bli tatt i bruk om kort tid. Basert på bibliotekenes behov har vi nå et prosjekt som skal gjøre det mye enklere å finne kvalitetskontrollerte nettressurser, der SPARQL er en helt sentral komponent. Etersom mange felter i IT-industrien har egne spørrespråk tenker vi oss at SPARQL vil bli en sentral komponent når forskjellige systemer skal snakke med hverandre.

Det er nettopp der SPARQL skiller seg fra tidligere teknologier. For det første bygger SPARQL på et allerede fungerende globalt informasjonssystem, nemlig webben. For det andre er språket lagd for å brukes på tvers av svært forskjellige datakilder.

Folk flest vil ikke se mye til SPARQL i nær framtid, ettersom språket er nokså innviklet. Det er likevel enklere å formulere SPARQL enn å programmere, og dermed vil ting som tidligere krevde kostbar og vanskelig programmering nå bli billigere og enklere. Dette betyr at teknologien er mer tilgjengelig og at flere vil benytte seg av den, noe som vi håper vil føre til en stor oppblomstring av nyttige ting fra grasrota, som var slik webben vokste i starten.

I tillegg til grasrot-vekst satser Computas AS stort på et voksende kommersielt marked, der vi forventer verdiskaping langt utover det dagens web kan gi. Flere store multinasjonale selskaper har

tilsvarende satsninger, noe en rekke firmaer uttrykker i forbindelse med pressemeldingen til World Wide Web Consortium [6] i sakens anledning.

- [1] <http://www.w3.org/2007/12/sparql-pressrelease>
- [2] <http://www.w3.org/2001/sw/>
- [3] <http://dbpedia.org/>
- [4] <http://geonames.org/>
- [5] <http://www.computas.com/>
- [6] <http://www.w3.org/2007/12/sparql-testimonial>

For mer informasjon kontakt:

Thomas Bech Pettersen

Direktør strategi og teknologi

Direkte: 67 83 11 59 | Mobil: 932 48 357

Epost: thomas.pettersen@computas.com

eller

Kjetil Kjernsmo

Senior Kunnskapsingeniør

Direkte: 6783 1136 | Mobil: 986 48 234

Epost: kjetil.kjernsmo@computas.com