

Oppdag ny kunnskap med WolframAlpha!



Det er grunn til å tro at WolframAlpha kan komme til å ta en del av de rene faktasøkene både fra Google og Wikipedia. For oss nordmenn og ikke minst bibliotekarer blir det selvsagt spennende å se hvordan WolframAlpha er tilpasset norsk språk og norske fakta.



Av Tore Brattli, førstelektor i dokumentasjonsvitenskap ved Universitetet i Tromsø (UiT)



I datamaskinens barndom var det stor tro på at den nye oppfinnelsen snart skulle kunne svare oss på avanserte spørsmål om alt fra kjernefysikk til meningen med livet. I dag vet vi bedre. Selv Google kan best sammenlignes

med en hund som er trent opp til å hente inn avisen hver morgen. Et enkelt gjøremål der den ikke trenger å forstå seg på innholdet i avisen.

Men nå lanseres det en kunnskapsmaskin som selv regner ut svaret, i stedet for å

returnere dokumenter som tilfeldigvis ligner på spørsmålet ditt. En maskin som skal kunne gi gode svar på spørsmål som ingen tidligere har stilt og som etter sigende skal inneholde nesten like mye fakta som i alle fagbibliotek til sammen.



Selv Google kan best sammenlignes med en hund som er trent opp til å hente inn avisen hver morgen. Et enkelt gjøremål der den ikke trenger å forstå seg på innholdet i avisen.

Superavansert kalkulator

WolframAlpha (<http://www.wolframalpha.com/>) er en helt ny type søkemotor spesielt beregnet på søk etter faktainformasjon. Egentlig er det ikke en søkemotor i klassisk forstand, men mer en slags superavansert

kalkulator som kan regne ut det aller meste med utgangspunkt i data, formler og oppskrifter. WolframAlpha (WA) kan gi deg svar på spørsmål som aldri før verken har blitt stilt eller besvart. Dette i motsetning til søketjenester som f.eks. BIBSYS

eller Google, som utelukkende returnerer dokumenter med nedtegnet kunnskap som forhåpentligvis kan besvare brukerens spørsmål eller hjelpe dem på veien.

Noen typer informasjon hentes opplagt best ut fra

«Målet har ifølge sjefsutvikler Stephen Wolfram vært å gjøre all vår sivilisasjons systematiske kunnskap beregnbar (!).»

tradisjonelle dokumenter som bøker, artikler, oppslagsverk eller nettsider. For andre typer informasjon er slike kilder mindre egnet. Et banalt eksempel er multiplikasjon. Det er bedre å kunne multiplisere enn å slå opp i en tabell med svar på alle mulige slags multiplikasjonsstykker. Et mer avansert eksempel kan være tidspunkt for alle totale solformørkelser i det 22 århundre sett fra Lindesnes. Et spørsmål som det er usannsynlig å finne svar på i Google og som det er vanskelig å regne ut selv for folk flest, men som vil være enkelt for WA.

For spørsmål om meningen med livet eller tips om ny sofa som kan matche den brune rokokkostolen er det nok bedre å bruke BIBSYS/Google. WA er først og fremst for faktasøk og målet har ifølge sjefsutvikler Stephen Wolfram vært å gjøre all vår sivilisasjons systematiske kunnskap beregnbar (!).

VG og New York Times

WA har utrolig nok vært en godt bevart hemmelighet frem til første artikkel ble publisert på Wolframs egen blogg (<http://blog.wolfram.com/>) 5. mars 2009. I den første tiden fikk WA mest oppmerksomhet i blogger, men også enkelte aviser som VG og The New York Times var tidlig ute med artikler.

Etter Wolframs demonstrasjon og webcast på Harvard universitetet 28. april, har oppmerksomheten økt betraktelig i alle medier. De fleste av dem som har fått demonstrert WA virker svært så entusiastiske for å si det mildt.

Bakgrunnen for WA

Og det er grunn til å ha store forhåpninger til Wolfram Alpha. Tjenesten er utviklet av firmaet Wolfram Research, mest kjent

for den matematiske programvaren Mathematica og sin genierklærte grunnlegger Stephen Wolfram. For tiden jobber 250 ansatte bare med WA. Wolfram skrev høyt siterte (441 i ISI-WoK) artikler innen partikkelfysikk allerede som 17-åring og fikk Ph.D. da han var 20. På 90-tallet forsket Wolfram mye på celleautomater (en slags primitive datamaskiner), noe som i 2002 resulterte i den relativt kontroversielle boka "A new kind of science" (NKS). Ideen her er at selv de mest komplekse ting kan beskrives ved hjelp av enkle regler. Som han selv sier om forutsetningene for WA:

«I had two crucial ingredients: Mathematica and NKS. With Mathematica, I had a symbolic language to represent anything—as well as the algorithmic power to do any kind of computation. And with NKS, I had a paradigm for understanding how all sorts of complexity could arise from simple rules.»

Som en kuriositet kan nevnes at Wolfram Research også er matematiske konsulenter for TV-serien Numb3rs.

Mer om hva WA er for noe og hvordan den fungerer

For brukeren består WA av et lekkert brukergrensesnitt med et enkelt søkefelt i kjent Google-stil. Søkespråket kan for mange formål også være likt. Et enkelt søk kan bestå av kun å hente fram data (f.eks. høyden på Mt. Everest), mens et mer avansert søk (f.eks. solnedgang 15. sept. samme sted) kan kreve både data og bruk av formler. Det eneste brukeren trenger å gjøre er å skrive inn spørsmålet, mens WA tar seg av data, formler, utregninger og selvsagt svaret. Mange enkle spørsmål har enkle svar selv om de er atskilt av

kompliserte utregninger.

Det er også grunn til å tro at WA også kan brukes som en avansert symbolsk kalkulator som kan løse ligninger, forenkle matematiske uttrykk, plote grafer, osv. WA består av 4 deler:

- **Dataforvaring:** Innsamling, vasking og organisering av data fra "statistiske sentralbyråer", oppslagsverk, håndbøker o.l. Foreløpig antydes det mer enn 20.000.000.000.000 "fakta". WA legger mye prestisje i å ha høy datakvalitet. Dette har vist seg å være ei stor utfordring siden kontroller har avslørt varierende kvalitet også fra anerkjente kilder. Brukere kan bidra med egne data, men det vil ikke være fritt fram som i Wikipedia. WA ønsker å sjekke alt som legges inn og kilder vil refereres i den grad det er mulig.
- **Beregningsdelen:** Består av formler, algoritmer (oppskrifter), modeller o.l. Selv om det selvsagt finnes et enormt antall slike formler, ser det heldigvis ut til å være et endelig håndterbart antall. I tilfeller der det eksisterer ulike oppfatninger (f.eks. ulike universmodeller) skal WA gi uttrykk for dette og gi brukeren mulighet til å velge ønsket modell.
- **Lingvistisk analyse av spørninger:** Det har vist seg enklere enn først antatt å tolke spørninger i naturlig språk. Tolkingen består i hovedsak av å knytte spørningen til et endelig antall entitets-typer, formler o.l. i WA, noe som er mye enklere enn å forstå tekst generelt.
- **Presentasjon:** For mange spørsmål finnes det nesten ubegrenset med mer eller mindre relevant informasjon som kan beregnes. Utfordringen har vært og automatisk kunne trekke ut det viktigste og presentere det på

en forståelig måte. Mange svar vil inkludere menyer for alternative presentasjoner og lenker til videresøk, bl.a. Wikipedia.

Allerede ved oppstarten dekker WA rundt 90 prosent av innholdet i et typisk reference library innen viktige disipliner som realfag, medisin, samfunnsfag og teknologi. I tillegg kan også brukerne bidra med sine egne data og formler, både som en del av spørringen eller ved å få dem inkludert i WA.

Noen vil kanskje få assosiasjoner til HAL 9000 og kunstig intelligens, men WA er ikke noen av delene. En bedre sammenligning er kanskje data fra Statistisk sentralbyrå, Meteorologisk institutt, Statens kartverk og en mengde andre institusjoner, sammen med alle mulige formler, oppskrifter og modeller. Til sammenligning er SSB er godt i gang med å kunne tilby mer fleksible presentasjonsmuligheter, men det er i hovedsak utdrag av ferdig utregnet statistikk og ingen verktøy for selv å kunne regne ut svar på egne spørsmål. Det WA ser ut til å ha klart er nettopp det å få alle slags data knyttet mot relevante formler og oppskrifter på en måte som gjør det enkelt å regne ut "alt" som ønskes.

Eksempler på WA-søk

Med forbehold om dekning innen norske forhold spesielt og bruk av norsk språk, kan dette være eksempler på søk i WA:

- Oslo = gir faktainformasjon

om byen

- weather Oslo 1990 = gir værstatistikk for Oslo for året 1990
- weather in Oslo when Jens Stoltenberg was born = 0,2°C i snitt, litt sol men for det meste overskyet
- 5 mol H₂SO₄ = bl.a. hva du trenger for å lage denne mengden svovelsyre
- jernbanelengde versus BNP = liste over land der Norge ser ut til å ligge på 42. plass
- rød + Gul = Oransje, i tillegg til bølglengde, HTML-kode o.l.
- president i Brasil i 1922 = Epitácio da Silva Pessoa og Artur da Silva Bernardes
- distance earth moon 21.07.2009 = 357464 km (varierer tross alt med 42592 km)
- 343g gull = verdi, volum, varmekapasitet, osv.
- Internet users in Europe = antall, fordelt på land, økning per år, osv.
- LDL 7,5 male age 40 = nivå av kolesterol satt i sammenheng
- lån 2.500.000 kr 3,25 % 20 år = renter og avdrag på lån
- D sharp minor = bl.a. ei lydfil som kan avspilles
- height Mt Everest / length Golden Gate bridge = 3,23273657
- forventet lengde på ekteskap i forhold til alder ved inngåelse = ???
- vitenskapelige ansatte per avlagte PhD ved norske universiteter = ???
- 2 glass appelsinjuice ei skive cheddarost = bl.a. næringsinnhold sammenlignet

- Forventet levetid for 50-årige kvinner fra Troms som røyker 20 sigaretter daglig og har BMI 35 = ???

Tilgang og abonnement

WA er i utgangspunktet gratis, men det vil komme tilbud om betalingstjenester der det f.eks. blir mulig å laste inn egne data, formler o.l. til mer spesialiserte formål. Litt reklame må man også regne med, men ikke mer enn det som i dag er vanlig i f. eks. Google.

Det vil være fleksible muligheter til å kommunisere med WA på ulike nivåer fra andre nettsider eller dataprogrammer ved bruk av API-er og svarene vil være tilgjengelig i flere formater inkl. PDF og Mathematica notebook (for mer avanserte beregninger).

Konkurrent til Google?

Flere har allerede spekulert i om WA kan bli en konkurrent til Google, men tjenestene har nok relativt ulike bruksområder. Likevel er det grunn til å tro at WA kan komme til å ta en del av de rene faktasøkene både fra Google og Wikipedia.

For oss nordmenn og ikke minst bibliotekarer blir det selvsagt spennende å se hvordan WA er tilpasset norsk språk og norske fakta. Uansett ser det ut til å være en utrolig interessant søketjeneste som åpner 29. mai 2009. (Bok og Biblioteks deadline var før denne datoen.)

«Noen vil kanskje få assosiasjoner til HAL 9000 og kunstig intelligens, men WA er ikke noen av delene.»