

Söknotom

Advanced Recommendations
Music: with smart

Find Results In:

Advanced Search | Refinements | Search Tips

11,000 Results (0.06 seconds)

smart - Click here for a list of intern...

See reviewed sites in: Television Cartoon
Hip-Hop Artists D-F | Television Cart...

1. CERT Advisory CA-98.01 "Smart" IP Denial-of-Service Attacks

Home | What's New | FAQ | Site Content | Contact Us

Other Resources |

URL: www.cert.org/advisories

[Translate](#)

2. <http://www.powertech.com>

Current top ten smart engines

Registered at Home AS...

URL: www.powertech.com

[Translate](#)

3. [Web Smart Club](#)

The Ultimate Web Smart Club

URL: websmartclub.org - Last...

[Translate](#)

[Smart the Web!](#)

Welcome to the WebSmarter!

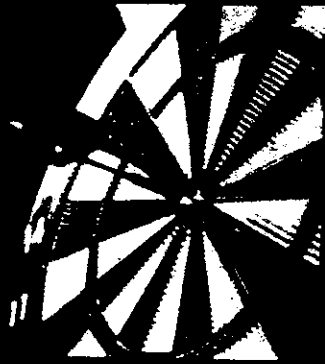
URL: websmarter.demall.net - L...

[Translate](#)

[Smart Off!](#)

Officially Anti-Smart Page

URL: www.paul.com/anti-smart



ernas krig

Davids kamp mot Goliat

Altavistas starka ställning bland svenska Internetanvändare håller i sig, men på hela Internet har de tappat mark. Altavistas huvudägare CMGI planerar att satsa cirka en miljard kronor på marknadsföring av Altavista. De vill att Altavista ska ta över ledartröjan i söktjänstkriget och bli en hörnsten i deras webbimperium. Men bakom hörnet lurar söktjänsternas nya komet Google. Med en unik rankingsteknik i bakfickan har de sakta men säkert fått fler och fler anhängare. Men det är en Davids kamp mot Goliat.

Yahoo, som är nätets populäraste söktjänst, har som många andra söktjänster följt med i portaliseringsvågen och har ett långt försprång framför andra söktjänster. Försprånget finns inte minst gentemot Altavista, som satsat i stort sett helt och hållet på själva söktjänsten. I grunden är dock Yahoo en helt vanlig webbkatalog med verkliga människor som redaktörer. Yahoo har ingen egenproducerad sökmotor.

Men Yahoo är inte bara en starkt portaliserad webbkatalog. Hittar den inte tillfredställande svar i webbkatalogen så hoppar den över och söker i ett index som någon annan sökmotor skapat. Sökningen sker med hjälp av en spindel. Vilket index den söker i har skiftat genom åren.

Redan 1996 valde Yahoo Altavistas index, då Altavista är den mest besökta sökmotorn bland svenska användare. Endast portalerna Msn.com, Spray och Passagen är mer besökta enligt statistik från Alexa. Yahoo.com hamnar på sjätte plats. Anmärkningsvärt är att den svens-

ka versionen av Yahoo, det vill säga Yahoo.se, kommer först på 42:a plats.

Altavista känns igen

Det finns ytterligare saker som talar för att Altavista har ett närmast orubbligt inflytande över svenska Internetanvändare. I en undersökning som Temo gjorde i juni i år svarade 69 procent av de tillfrågade att de kände till sökmotorn Altavista. Konkurrenten Yahoo kände bara hälften av de tillfrågade till, medan ingen nämnde Google. Google är annars söktjänsternas fräckaste uppstickare på senare tid.

Med tanke på Yahoos odiskutabla popularitet är det förmodligen den finaste belöningen en sökmotor kan få när den får sälja sitt index till Yahoo. Men Altavistas belöning varade bara i två år. Inktomi valdes i maj 1998 som ny samarbetspartner till Yahoo.

Altavistas ställning på hela Internet är inte lika stark som i Sverige. Enligt Media Matrix hamnar de på åttonde plats bland söktjänster och enligt Alexas

statistik halkar de kontinuerligt efter i listan. Det kanske till och med är så att Altavistas portaliseringshysteri blivit ett problem?

– När jag söker med Altavista kommer det jag är intresserad av ofta långt ner på listan. Google blir däremot bara bättre och bättre. Det är nog vår smartaste konkurrent, sa Peter Nordin till tidskriften Ny Teknik i augusti 2000. Han är docent på Chalmers och grundare av söktjänsten Leika, som bygger på artificiell intelligens.

Relevanta svar

Låt oss ta ett konkret exempel på problemet med att leverera relevanta svar. Låt säga att du är intresserad av den omtyckta tjänsten Direct Search, som länkar till information på den osynliga webben. Bara ordet "direct" ger drygt 2,5 miljoner träffar i Google och ordet "search" drygt 2,8 miljoner. Tack vare Googles utmärkta rankingsteknik hamnar Gary Prices webbtjänst Direct Search som första träff när man söker på

Altavista - Welcome Netscape

altavista:  www.GetAWebAddressNow.com [CLICK HERE](#)

Comparison Shop Channels Rewards Email & Tools

[Sign Up!](#) | [Altavista Members Sign In](#) | [Member Center](#) | [My Altavista](#)

FIND HOT ARTIST SITES: [Metallica](#) | [Jay-Z](#) | [The Beatles](#)

any language Search

Today In:

| | | | |
|--|---|--|--|
| Shopping Holiday Shopping Guide Win a Trip to D.C. \$10,000 Handhelds Searches 1. Palm Vlx 2. Casio EM 500 3. Pison Series QMX | Entertainment Sharon Stone: Fat Photo? Caine Knighted by Queen Hot Model Searches 1. Claudia Schiffer 2. Giselle Bundchen 3. Naomi Campbell | News Crucial Court Calls Loom 7 New Web Surfboxes OK'd Today's News Searches 1. Clinton Arrives in Vietnam 2. Nasdaq Sinks 4 Percent 3. Barak Long Fight Ahead | Real Estate Buy or Sell a Home Apply for a Mortgage Real Estate Searches 1. Get an Apartment 2. Find Senior Housing 3. Find Real Estate Agents |
|--|---|--|--|

Altavista Resources

Find It: [Autos](#) | [Real Estate](#) | [Planet Project™](#) | [Legal Help](#) | [Fashion & Beauty](#) | [Joke of the Day](#)

Services: [Health](#) | [Horoscopes](#) | [Movie Showtimes](#) | [MyPrimeTime](#) | [Find a Lawyer](#) | [Small Business](#)

Tools: [Domain Name Registration](#) | [Media Player](#) | [Submit a Site](#) | [PC HelpDesk](#) | [Downloads](#) | [Travel](#)

Find by Category

Altavistas ställning är fortfarande stark i Sverige. Frågan är om det beror på att folk går dit av gammal vana, eller om de verkligen tycker att det är den bästa söktjänsten.

båda orden. Av sammanlagt 1,3 miljoner träffar dessutom.

Gör samma sökning på "direct search" i Altavista, och Gary Prices webbtjänst finns först som 14:e träff. Dessutom är träffen inte på ingångssidan, utan det är ett annat dokument som hittats. Det är ett tydligt exempel på Altavistas brister i fråga om att leverera relevanta svar.

– Altavista håller inte längre den kvaliteten som de tidigare har gjort och deras grundläggande teknologi motsvarar inte de krav vi har, sa Magnus Gagnelid, projektansvarig för Evreka.com, till DN när de valde Fasts index i februari 2000.

Altavistas tid som leverantör av webbindex verkar vara till ända. Det är nu länge sedan de blev överkörda av Inktomi och förlorade samarbetet med Yahoo. Nu har de också förlorat Evreka.com som kund.

Störst och först

Altavista var ute på nätet redan i december 1995 och kunde snabbt visa på det största indexet. Framför allt i

Internethysterins Sverige fick de många spaltmeter i diverse tidningar i samma veva som allt fler började koppla upp sig. Altavistas starka ställning är ett tydligt tecken på hur viktigt det är att vara ute i rätt tid, något som alla e-handelsföretag inte tycks ha förstått när de satsar miljon på miljon innan ens svenskarna har fått förtroende för säker handel via nätet.

Men tiden håller på att äta upp Altavista, som möter mördande konkurrens både på portal- och sökmotormarknaden. Kan de inte hitta en lösning på rankningsproblemen eller med någon magisk formel lyckas i portalkriget, så lär de med all säkerhet hamna lika mycket i skymundan som pionjären Webcrawler, en relik från Internets barndom då ingen ens talade om webbtäckning.

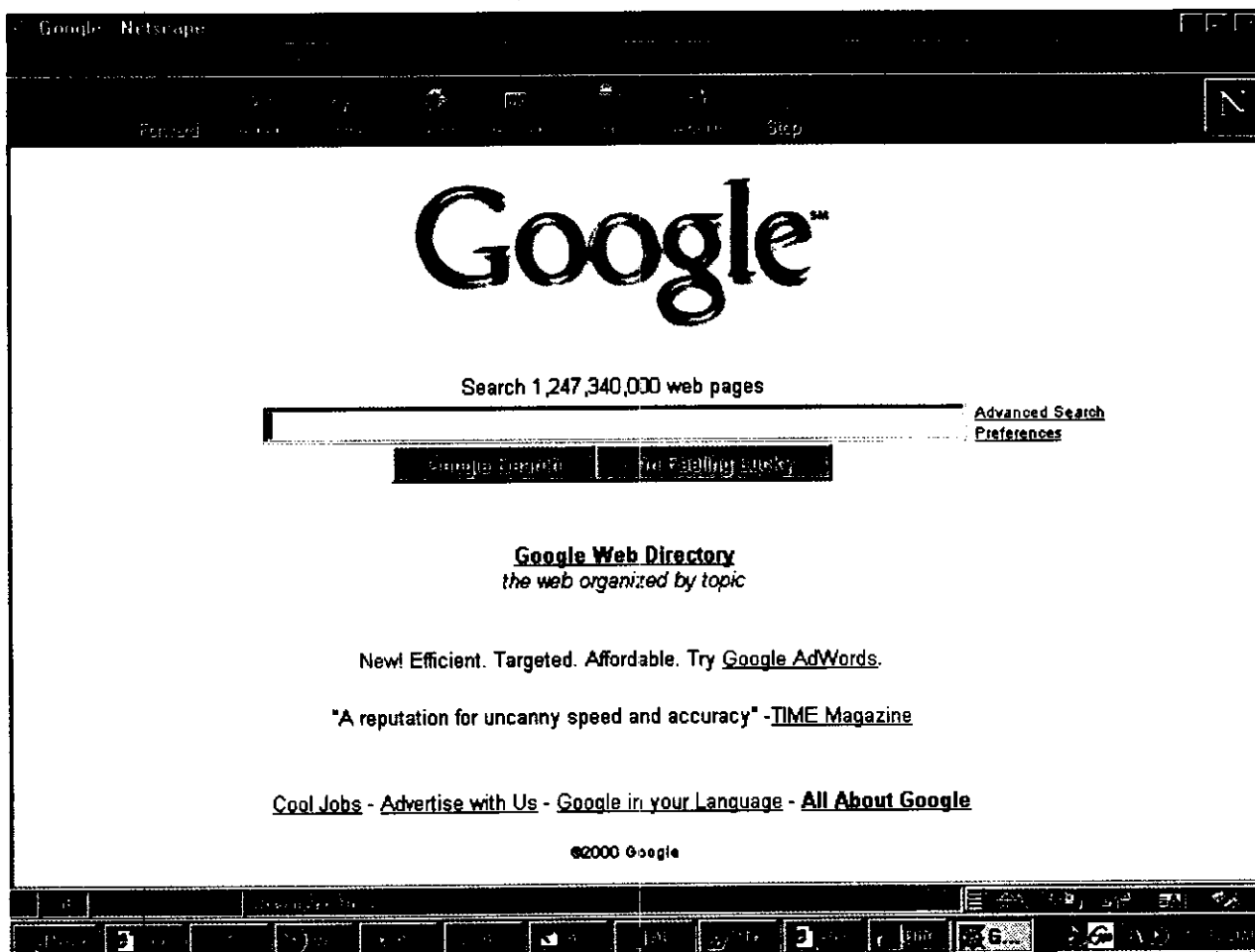
Altavista har under en längre tid också haft ekonomiska problem. Det är visserligen inget de numera är ensamma om bland dotcom-bolagen. Redan 1999 sålde Compaq företaget till riskkapital-

bolaget CMGI och man planerade för en börsnotering som nu har skjutits på framtiden. I stället har man under den gångna hösten skurit ned antalet anställda med 25 procent, och talesmän från Altavista hävdar att man kommer att renodla verksamheten till sökfunktioner. Man tycks med andra ord ha gett upp portalambitionerna.

Snabbast på Alphaserverar

Altavista utvecklades en gång i tiden av dataföretaget Digital med syftet att bland annat visa kraften i deras Alphaserverar. Mannen som hittade på alltihop, Paul Flaherty, märkte i april 1995 att version 7 av Oracles databas gick hundra gånger snabbare på deras Alphaserver än på någon annan maskin på marknaden. Ingen annan söktjänst kunde visa på lika snabba sökningar och lika stort index som Altavista. Det var Altavista som sådde fröet till det som senare kom att kallas "The Size War", det vill säga webbtäckningskriget.

Men det räcker inte enbart med stor



Google är en liten uppstickare som på många sätt överträffar mångmiljardindustrin Altavista. Enkelt, rent, snabbt och bra rankade träffar är konceptet.

hårdvarukapacitet och bra webbtäckning. Tillväxten av antalet sidor på nätet har nått sådana dimensioner att det helt enkelt är ett dåligt skämt om en sökmotor påstår att den täcker hela Internet. Bland annat av den enkla anledningen att det är omöjligt att bevisa att man täcker hela Internet.

Att bevisa motsatsen är desto lättare. En söktjänst måste till syvende och sist kunna presentera relevanta svar för att behålla sina besökare.

Det är precis av denna anledning som nätets stora uppstickare Google har skördat så många framgångar under det gångna året. Som PC World uttryckte det i septembernumret 2000:

"Google lovar att presentera webbens mest relevanta svar – och de levererar det."

Enkelt och avskalat

Googles tanke från allra första början har varit en så kallad "Less is more"-strategi. I en allmänt hysterisk tid av portalisering dök Google upp med sitt

Cranfield-metoden

"Recall" och "precision" är vanliga termer inom informationsvetenskap som används när man utvärderar informationssökningssystem. Termerna används i en känd utvärderingsmodell av Cranfield, som helt enkelt brukar kallas Cranfield-metoden. "Precision" mäter antalet relevanta återfunna dokument i förhållande till hela antalet återfunna dokument. Med detta kan man ange hur mycket ovidkommande information, det vill säga brus, ett sökresultat innehåller. Som matematisk formel kan den uttryckas så här:

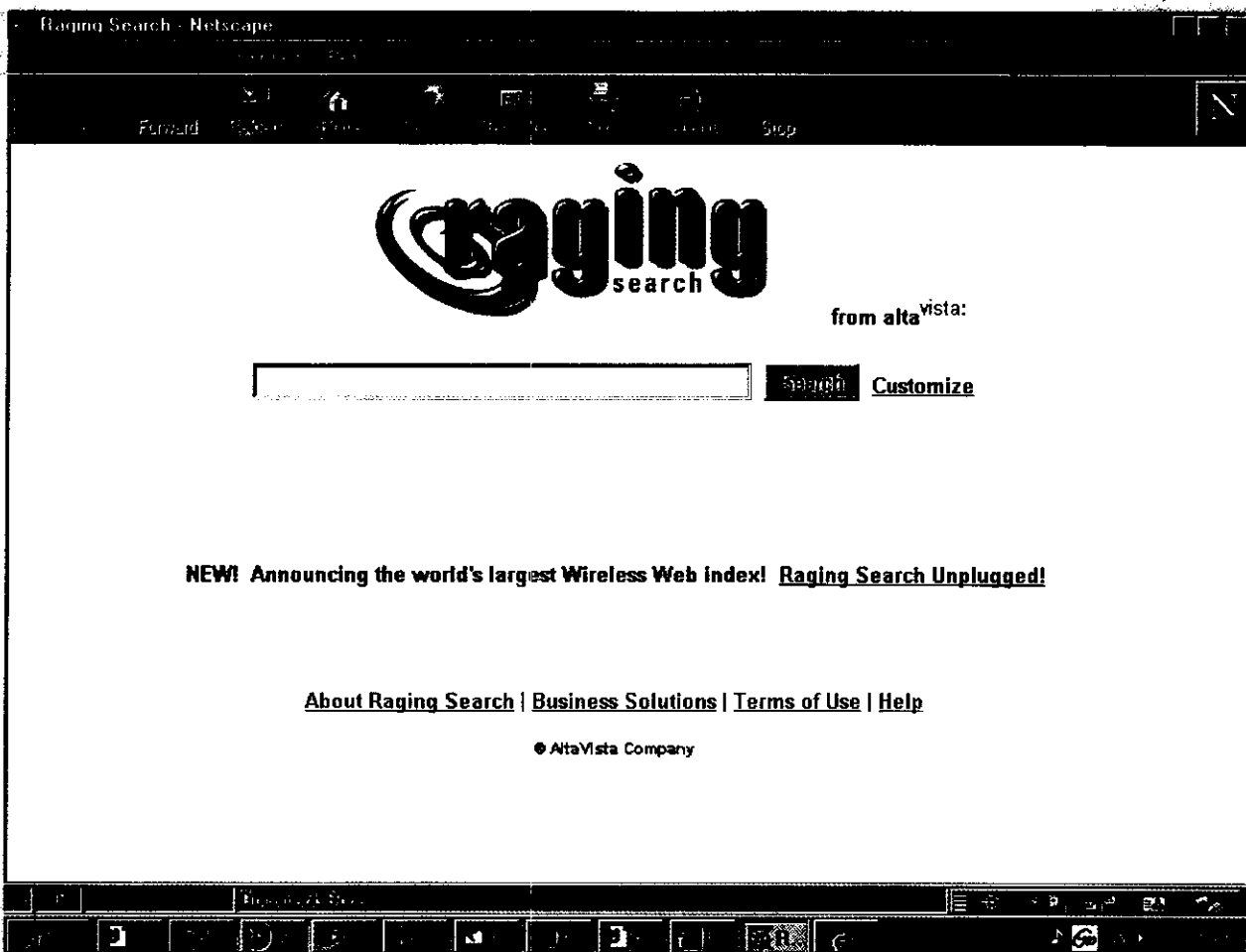
$$p = \frac{\text{relevanta återfunna dokument}}{\text{återfunna dokument totalt}} \times 100\%$$

"Recall", eller anträffningsgrad, mäter antalet relevanta återfunna dokument i förhållande till det totala antalet relevanta dokument i samlingen. Det anger hur uttömmande sökningen är. Som matematisk formel kan det uttryckas så här:

$$r = \frac{\text{relevanta återfunna dokument}}{\text{relevanta dokument i samlingen}} \times 100\%$$

Både "precision" och "recall" är svåra mätmetoder för söktjänster på webben med tanke på de enorma mängder webbsidor som finns i indexen.

"Precision" går hjälpligt att använda när få träffar presenteras "recall" går över huvudtaget inte att mäta, eftersom det är omöjligt att definiera antalet relevanta dokument i en sökmotors index.



Altavistas lansering av den nya söktjänsten Raging.com är en tydlig markering av att de nu tar upp kampen med Google. En av idéerna är precis som i Googles fall att det ska vara en ren sida utan en massa kringtjänster – bara sökning.

till synes opretentiösa gränssnitt präglad av en stram formgivning, med så lite knappar och kringtjänster som möjligt. Ju färre val, desto bättre.

Eller som Larry Page, den ene av Googles två grundare, beskriver det i en intervju i sommarnumret av tidskriften Online:

– Vi håller Google fritt från allt virrvarr. Det finns massor av forskning som visar att den tid det tar dig att göra

någonting är proportionellt mot antalet valmöjligheter du har. Om du går till en webbsida med en massa länkar, så tar det egentligen längre tid att göra en sökning där än om du söker i Google, för där finns inte lika många valmöjligheter.

Googles fokusering är tydligt inriktad på sökning och inte på portalisering, en strategi som allt mer börjat ge resultat. Altavistas lansering av Raging.com i maj i år är ett tydligt försök att ta upp kam-

pen med "Less is more"-strategier som Google och Fasts Alltheweb.com. Med fina ord som "högkvalitativt index på 350 miljoner sidor" och "ny patenterad relevansalgoritm" och en "relevanstest från ZD Labs" som råg i ryggen tog Altavista nya tag.

Ööverträffad ranking

Men huvudingrediensen i framgångskonceptet Google är trots allt deras tek-

Sökspråk för Altavista och Google

| Sökmotor | Booleskt | Närhet | Trunkering | Case | Fält | Begränsningar | Stoppord | Sortering | Klustring |
|-----------------------------|----------------------|------------|------------|------|------------------------|---------------|----------|---------------------------------|---------------|
| Altavista enkel sökning | +,- | Fras | Ja* | No | title, URL, link, m fl | Språk | Ja | Relevans, citering | Ja, förval |
| Altavista avancerad sökning | and, or, and not, () | Fras, Near | Ja* | Yes | title, URL, link, m fl | Språk, datum | Nej | Relevans, om använt | Ja, ej förval |
| Google | -, OR, and | Fras | Nej | Yes | Link, related | Språk, domän | Ja | AskJeeves, RealNames, Relevance | Ja, förval |

Google söker alltid på sidor som innehåller alla ord i din sökning, så du behöver inte använda + framför orden. Men + kan användas om man vill tvinga systemet att söka på stopp-

ord. Med stoppord menas ord som är väldigt vanliga och ofta saknar betydelse i en sökning, till exempel the, and, of, not...

nik att ranka träffar. Idén bygger på det som citeringsdatabaser (referensdatabaser) sysslat med i årtal, nämligen att räkna hur många citeringar en artikel har. Google överförde detta till webben och började räkna antalet länkar till en viss webbsida, det vill säga länkpolaritet. Den rankingmetoden har dock en stor brist. Den är inte särskilt svår att spamma. Den går att manipulera genom att man skapar fejsidor med länkar till den webbsida man vill ska rankas högt.

Google tänkte ett steg längre och lade till kriteriet att de inlänkar som i sin tur kommer från andra sidor med många inlänkar rankas högre. Plötsligt var det frågan om länkvärde och inte bara popularitet.

– Det är länkpolaritet ni sysslar med och inget annat, sa Andrei Broder från Altavista, med visst avund, under en paneldebatt vid årets konferens Search Engines Meeting i Boston.

– Nej, det är länkvalditet, svarade Googles Larry Page som lyste av självförtroende. Vilket han gjorde med all rätt.

Förmodligen visste han redan då att han var på väg att bli tilldelad en av de finaste belöningarna. Två månader senare meddelades att Yahoo bytt samarbetspartner för sin sökmotor från Inktomi till just Google. I septembernumret av PC World valdes Google till

den bästa söktjänsten. Som om inte det var nog fick Google av tidskriften Wireds läsare utmärkelsen Readers Raves Award i oktober 2000 för att de är "den mest intelligenta agenten".

Datamining

Frötet till Google såddes på Stanfords universitetet 1995 när Larry Page träffade Sergey Brin, som redan bedrev forskning om "datamining". "Datamining" är en term inom informationssökning som handlar om att analysera stora mängder av data och försöka hitta relationer och samband som gör det möjligt att hantera dem. Larry började undersöka webben och samlade ihop stora mängder data. Sergey blev intresserad och tillsammans började de experimentera.

I hela tre år försökte de undvika (!) att starta ett företag eftersom de var ganska nöjda med livet på Stanford, även om de kände viss lockelse att starta ett dotbolag. Till slut blev dock behovet av ännu mer hårdvara för stort. 1998 lanserades Google med en startplåt på 100 000 amerikanska dollar som de lyckats få av Andy Bechtolsheim, en av grundarna till Sun. Och på den vägen är det.

Fler och fler sökare med erfarenhet har börjat upptäcka Googles förträffligheter och detta ofta efter att ha drunknat i de irrelevanta svar som till exempel Altavista ger. Altavista ger oftast fler svar

än nödvändigt. Frågan är då om det bara är en allmän känsla hos användarna eller om det finns logiska förklaringar till det.

Förvald operator

Något som är av väsentlig betydelse när man söker i en söktjänst är vilket förval som erbjuds vid sökning, till exempel om det är operatoren AND eller OR som är förvald. Om man söker på två ord betyder OR, i Altavistas fall, att dokument där minst ett av orden förekommer presenteras. Så är inte fallet i Google.

Förvalet i Google är i stället AND. För varje ord som läggs till i sökningen betyder det att ordet måste förekomma i dokumentet. Man får alltså inte som i Altavistas fall dokument där bara ett eller fler av orden förekommer.

– Vi föredrar "precision" istället för "recall". Man får färre träffar i svarslistan, men det är mer sannolikt att de är de dokument du ville ha, förklarar Larry Page från Google i intervjun i Online.

En erfaren användare märker ganska snabbt fördelen med AND som förval och fokuseringen på "precision" framför "recall". Det är lättare att ha kontroll över sin sökning.

Vill man komma runt detta i Altavista kan man i enkel sökning lägga till ett plus framför de ord man anser måste

Statistik är alltid statistik

Att all statistik måste tas med en nypa salt och mer ses som en vägledning är väl inget nytt. Särskilt gäller det den besöksstatistik som det hänvisas till i den här artikeln. Det är visserligen den bästa statistiken man kan få och den är ofta väldigt användbar, men som all annan statistik måste den analyseras.

Ta till exempel besöksstatistiken för Msn.com, som är den mest besökta webbplatsen bland svenska användare enligt Alexa. Vad man då bör veta är att alla användare av e-posttjänsten Hotmail som använder Hotmails "log out"-funktion dirigeras till just Msn.com.

I Alexas statistik kan man också se att över 80 procent av alla besökare har kommit dit via länkar från Hotmail. Hotmail, som är en av nätets populäraste e-posttjänster, drar alltså till Msn.com en mängd besökare som förmodligen aldrig hade en tanke på att surfa in dit.

Media Matrix, som är webbens äldsta värderingstjänst, presenterar varje månad statistik över de mest besökta

| Ratingservice | Media Matrix (oktober 2000) | | Netratings (juni 2000) | | Alexa (september 2000) | |
|---------------|--------------------------------|-----------|---------------------------|-----------|---------------------------|-----------------------|
| | Räckvidd | Placering | Räckvidd | Placering | Placering Sverige | Placering Alla länder |
| Altavista.com | 17,7% | 8 | 12,3% | 6 | 4 | 14 |
| Google | 7% | 14 | 3,8% | 12 | 39 | 23 |

Räckvidd betyder en uppskattning av antalet webbsurfare, i procent, som besökt en söktjänst under en viss månad. Eftersom webbsurfare kan ha besökt flera söktjänster kan de sammanlagda procenten från alla söktjänster överstiga 100. Media Matrix och Netratings bygger på detta. Alexa räknar sammanlagd trafik under en viss månad.

webbplatserna. De delar dock in dem i två olika grupper, nämligen så kallade "top web properties" och "top web sites". Digital property som inte nödvändigtvis finns tillgänglig via webben är också medräknad. "Web property" definieras som en samling av webbplatser ägt av ett företag. Lycos till exempel äger flera olika webbplatser. Förutom den egna portalen med samma namn äger de den svenska portalen Spray, sökmotorn HotBot och webbhotellservicen Tripod. Besöksstatistik från alla dessa tjänster ingår alltså i statistiken för lycos "web property".

Följaktligen betyder "web sites" statistik för varje enskild webbplats, oavsett om olika webbplatser har samma ägare.

Besöksstatistiken för söktjänster som har portaliserats innebär att även all surfning som inte har med sökning att göra, det vill säga att läsa e-post, läsa värdet etcetera ger ökad trafik. Det innebär att bara en del av statistiken för dessa portaliserade söktjänster gäller sökning.

Den sammanlagda trafiken till en portal kan därför vara större än till en sökmotor, men samma sökmotor kan i sin tur ha fler besökare än samma portal.

1. Världens Bästa Mamma!

...
URL: www.torget.se/users/z/Zara/mor/intr.htm - Last modified on: 10-Feb-2000
[More pages from this site](#)

Bild 1.

1. Lycos Sökning: champagne mousserande vin Frankrike druvor

Hjälp Service. Sökning: Avancerad Sökning. Lycos Homepage > Sök > Resultat. Nyheter. Annonsering på Lycos Lägga till min hemsida Kontakt Home On The.
URL: www.lycos.se/cgi-bin/pursult?cat=graphic...ror&ignore=lang - Last modified on: 10-Feb-2000 - In Swedish
[More pages from this site](#)

Bild 2.

VinNytt

... kring vinet i Paris och i Frankrike från Vinprovningar Britt Karlsson. Det ... på vinbarer och restauranger, artiklar om vin, druvor och kombinationer mat och vin ...
ourworld.compuserve.com/homepages/vinnytt/homepage.htm - 5k - Cached - [Similar pages](#)

Bild 3.

förekomma i dokumentet. Vad Altavista också tittar på, innan den söker med OR mellan två eller fler angivna ord, är om det finns en fras med samma ord.

- En sökning på: *världens bästa mamma* gav i Altavista resultatet som du kan se i bild 1.
- En sökning på: *+frankrike +vin* gav i Altavista resultatet du kan se i bild 2.
- En sökning på: *frankrike vin* gav i Google resultatet du kan se i bild 3.

Relevanstest

Men om vi försöker att bortse från den subjektiva känslan av att en söktjänst presenterar bra eller dåliga svar och i stället försöker oss på en objektiv analys, så kanske vi kommer fram till helt andra slutsatser. Eller?

I det tidigare nämnda ZD Labs relevanstest, där fem stora söktjänster jämfördes, fick Altavista högst sammanlagd poäng (Se tabell 1). Fem olika typer av sökningar inom fem olika ämnesområden utfördes. De fem olika typerna av sökningar var:

1. naturligt språk
2. ett sökord
3. flera sökord
4. komplex sökning med flera sökord
5. sökning efter specifik webbplats

De tio första träffarna analyserades och gavs olika många poäng beroende på relevans.

Det visade sig att Altavista var bäst på

sökningar med flera sökord och komplexa sökningar med flera sökord. Google var överlägset bäst på ett sökord och sökningar efter specifik webbplats. Alla söktjänster fick låga poäng på sökningar med naturligt språk.

Googles rankingsteknik har sina skönhetsfläckar, vilket till och med Googles Sergey Brin erkänner i en intervju i Technology Review från november 2000:

– Till exempel: medan tekniken fungerar utmärkt för att mäta det relativa

värdet av en specifik webbsida, klarar den inte nödvändigtvis av att mäta om den är relevant: ger verkligen den här webbsidan svaret på min fråga?

Datormagazin gjorde ett förenklat relevanstest, liknande ZD Labs, för att se om det låg någonting i deras slutsatser (Se tabell 2).

I motsats till i ZD Labs test visade sig Google vara överlägset Altavista, även på tre sökord. I ZD Labs test bestod sökningen på flera sökord visserligen av fyra till sex sökord. Å andra sidan finns det undersökningar som visar att användare sällan söker med flera sökord.

Google fungerar särskilt bra på ett sökord och specifika webbplatser enligt båda testerna. Datormagazins test visade också att Google var extra skicklig på att presentera relevanta svar i de två första träffarna, vilket Altavista inte var.

Trunkering

Men Altavista har väl också sina förtjänster, förutom att de klarar komplexa sökningar bra?

Om man tittar på vilka avancerade sökmotorer erbjuder ligger Google lite i lä. I Altavista kan man söka med operatoren NEAR, vilket de är helt ensamma om bland de stora söktjänsterna. Det är obegripligt att inte fler söktjänster erbjuder denna avancerade sökmöjlighet.

När man söker i stora och ostrukturerade informationsmängder är operatoren

Relevanstest

Tabell 1. ZD Labs relevanstest

| Sökmotor | Altavista | Google |
|----------------------------------|-----------|--------|
| Naturligt språk | 34.5 | 33.5 |
| Ett sökord | 57.5 | 73 |
| Flera sökord | 63.5 | 32 |
| Komplicerad sökning på flera ord | 70.5 | 65.5 |
| Specifik webbplats | 67.5 | 79 |
| Totalt | 293.5 | 283 |

Tabell 2. Datormagazins relevanstest

| Sökmotor | Altavista | Google |
|--------------------|-----------|--------|
| Ett sökord | 36% | 70% |
| Tre sökord | 40% | 92% |
| Specifik webbplats | 63% | 100% |
| Genomsnitt | 46% | 87% |

Datormagazins relevanstest bygger till stora delar på samma metod som ZD Labs, men det är mycket mer förenklat. Tre olika typer av frågor har ställts i tre olika ämnen: mat, idrott och medicin.

Sökmotorer:

1. Ett sökord, till exempel ekorre.
2. Tre sökord, till exempel göteborg skärgård hamnar.
3. Specifik webbplats, till exempel svenska dagbladet.

Naturligt språk och komplexa frågor med flera ord har inte tagits med. De tio första svaren i varje sökning har poängsatts. Poängsatsen är omräknad i procent. I sökningen efter en specifik webbplats har 100 procent utdelats om webbplatsen funnits som första träff, 90 procent för andra träff, 80 procent för tredje träff, och så vidare. En poäng delas ut för relevant träff. Med relevant träff menas inte bara ordmatchning utan också att webbsidan ger information om ämnet i frågan. Två extra poäng om första träffen är relevant och en extra poäng om andra träffen är relevant. En halv poäng delas ut om det bara finns en länk som hänvisar till relevant information, och noll poäng om det inte är ett relevant svar eller om det är en spegelsida, en dubblett eller en död länk.

NEAR i de flesta fall mycket mer användbar än booleska operatörer som AND, OR och NOT. NEAR betyder, i Altavistas fall, att orden kan förekomma inom tio ords avstånd och i godtycklig ordning.

I motsats till Google tillåter Altavista också trunkering. Att Google inte använder automatisk trunkering eller "stemming" som förval ska de ha all heder av, men att trunkering inte ens finns som tillval är märkligt.

Så här förklarar Larry Page saken i tidskriften Online:

– Vi var väldigt tveksamma till fördelarna med att tillåta "stemming" i någonting så stort som webben. Man kan argumentera för att webben fördubblas ganska ofta – var åttonde månad ungefär – och ju större den blir, desto mindre bryr du dig om "recall" – alltså hur många dokument du får i träfflistan – men mer och mer om att det är rätt

dokument. Visst, både "stemming" och automatisk trunkering ökar antalet träffar, men med trunkering som tillval väljer man själv när det är användbart. Särskilt i komplexa sökningar, som kanske innehåller både booleska metoder och flera sökord och därför ger färre träffar, kan trunkering vara guld värt. Men Google har som sagt sin unika rankningsteknik som också ska vara omöjlig att spamma, en ren och tydlig ingångssida utan en massa kringtjänster och en cachefunktion där du kan se hur den aktuella webbsidan såg ut när den indexerades, och de har AND som förval, som vi tidigare nämnt fördelarna med.

Slutet för Altavista?

Är Altavistas tid som leverantör av webbindex slut? Google och Inktomi har ju för tillfället ett försprång. Eller kommer det att bli lättare för Altavista

att ta upp kampen nu när de har gett upp portaliseringen och fokuserar på sökning? Hur länge kommer Googles rankningsteknik att vara nätets smartaste? Kommer söktjänster som bygger på artificiell intelligens att ta över?

Så åter till frågan: Altavista OR Google? Svaret blir nog: Altavista AND Google.

Så länge Altavista är ensamma om sin NEAR-operator kommer inte jag personligen att överge Altavista, även om Google är mitt förstahandsval. Men det är alltid frågeställningen som bör styra valet av söktjänst.

Än är det dock långt kvar till den ultimata söktjänsten, trots att Google har tagit täten med sin unika rankningsteknik. Det gäller bara att få de oerfarna sökarna att upptäcka den. Den dagen kommer David med all säkerhet att besegra Goliat.

Lars Iselid

Ordlista för vetgiriga sökare

Artificiell intelligens (Artificial intelligence, eng.)

Program som efterliknar mänskligt tänkande.

Booleska operatörer (Boolean operators, eng.)

Benämning på ett avancerat sökspråk som består av följande tre operatörer: AND, OR och NOT. De ger möjlighet att inkludera och exkludera sökord. Så här fungerar operatörerna:

fotboll AND sverige – ger dokument som innehåller båda termerna

fotboll OR rugby – ger dokument som innehåller åtminstone ett av orden

fotboll NOT amerikansk – ger dokument som innehåller det första sökordet men inte det andra

I till exempel Altavista Avancerad sökning ska NOT-funktionen skrivas: **fotboll AND NOT amerikansk**

Citeringar (Citations, eng.)

En vetenskaplig artikel innehåller i slutet hänvisningar till källor, ofta andra artiklar, som man bygger sina slutsatser på. En artikel som många andra har citerat kan ofta bedömas vara mer värdefull än andra.

Inlänkar

En webbsidas inlänkar är de länkar som pekar till den sidan. I till exempel Altavista kan du kolla hur många som har länkat till en viss webbplats eller -sida genom att skriva *link:* och webbadressen.

En sökning på *link:www.altavista.com* gav 179.997 träffar och en sökning på *link:www.google.com* gav 53.136.

Rankning

I vilken ordning träffarna i en sökning presenteras. Används ofta i samlat begrepp tillsammans med relevans, nämligen relevansrankning.

Naturligt språk (Natural language, eng.)

Naturligt språk betyder sökningar som är utformade som en fråga, till exempel: *Where can I find pictures of libraries?* Prova att ställa frågan i söktjänsten Askjeeves.com, som är nätets ledande söktjänst för att hantera naturligt språk.

Relevans (Relevancy, eng.)

Att de träffar i en sökning som anses vara mest relevanta presenteras först. Varje söktjänst har olika kriterier för hur man bedömer träffarnas relevans. Google bygger sina relevansskriterier främst på länkvaliditet.

Spamma (Spamming, eng.)

I detta sammanhang: att manipulera en sökmotor genom att utforma en webbsida så att den rankas högt i en träfflista. Googles rankningsteknik PageRank anses vara spammsäker.

Stemming

Man tar bort prefix och suffix i ord och sammanför ord med samma "stam", för att göra dem sökbara tillsammans. En sökning på "gråt" ger även "gråta", "gråtarska", "gråter", "gråtmild".

Den vanligaste varianten av stemming är att sökning sker både på pluralis och singularis.

Trunkering (Truncation, eng.)

Möjlighet att söka på bara en del av ett visst ord. Vanligast är högertrunkering. En symbol, oftast *, anger att sökning ska göras på olika ordvarianter. Till exempel: "sjukhus*" ger även följande ord: "sjukhuset", "sjukhusentré", "sjukhusanläggning" och så vidare.

Vänstertrunkering

Innebär att en sökning på "*post" även ger "vaktpost", "dörrpost", "e-post" och så vidare.

Trunkering mitt i ett ord brukar kallas maskering och innebär att en sökning på "eri*on" ger "erixon", "ericsson", "eriksson", "erickson" och så vidare.