

## Online- och webbnyheter

Nr 2004:7 2004-12-30

### ***Rapport från Online Information 2004, London, 30 nov - 2 dec 2004***

Lars Klasén  
InfoData AB  
[lars.klasen@infodata.se](mailto:lars.klasen@infodata.se)

*Den 28:e upplagan av Online Information, Online 2004, var i de flesta avseenden väldigt lik förra årets tillställning. Nästan samma antal konferensdelegater, utställare och utställningsbesökare. Den var sig också lik såtillvida att den ganska väl speglade "uppmärksamhetsläget" i onlinebranschen, med "informationshantering" i förgrunden och "information" mera i bakgrunden - trots att det såklart är det senare området som inte bara motiverar utan också bär upp det hela, inte minst marknads-mässigt. Logiskt nog dominerades utställningsdelen ytmässigt av informationsaktörer (onlinetjänster, förlag etc), medan utställare inom informationshantering tog mycket mindre yta men till antalet ändå var i nästan samma storleksordning. Konferensdelen dominerades av informationshantering, med inslag även av webbsökning, medan onlinetjänster och databaser nästan inte alls berördes.*

*Ändå var det något speciellt med Online i år. Man kan uttrycka det som att Googles ande vilade över hela tillställningen, från konferens till utställning. Det berodde förstås på allt som hänt under det senaste året vad gäller webben, manifesterat i Google. Google fördes på tal i samtliga föredrag jag åhörde. Och att Google har ett starkt inflytande på hur tjänster och produkter utformas är ovedersägligt. Jag kallar det Googlemania ...*

*Rapporten, som är på 32 sidor, speglar också detta. Förutom "naturligt" i de sedvanliga beskrivningarna av utställare och föredrag så genom ett "extra" avsnitt om Googles enterprise-söksystem. Jag anser det motiverat eftersom Googles förehavanden är av stor betydelse för onlinebranschens utveckling.*

#### **Innehåll:**

- 1. Allmänt**
- 2. Konferensen**
- 3. Utställningen**
- 4. Utställare**
- 5. Google Search**
- 6. Föredrag**

### **1. Allmänt**

Det mesta var som vanligt på årets Online Information, det 28:e i ordningen. Det ägde rum under tre dagar i konferensanläggningen Olympia i London. Arrangör var VNU Exhibitions Europe (f d Imark Communications), där den tidigare arrangören Learned Information numera ingår. Den bestod av en konferensdel (med totalt ca 90 föredrag i tre sessioner) och en utställningsdel (i Olympias största utställningshall Olympia Grand Hall) som också erbjöd ett stort antal gratispresentationer mm. Dagen före evenemanget genomfördes ett antal hel- eller halvdags workshops. Antalet konferensdelegater var knappt 650, antalet utställare ca 250 och utställningsbesökare 10.600 - så även i dessa avseenden var det alltså ungefär som vanligt. (Mer information om Online 2004: [www.online-information.co.uk/](http://www.online-information.co.uk/))

## Online - en branschspegel?

Mer intressant är att evenemanget som vanligt också något så när speglade läget i onlinebranschen. Det innebar att den största uppmärksamheten riktades mot informationshantering (IA, CM, KM etc), men att evenemanget i sig bars upp av dess förutsättning, dvs *information*. Precis som "i verkligheten" alltså, där uppmärksamheten kring informationstjänster minskat i takt med att branschen stabiliserats och mognat, aktörerna blivit färre (men större) och information som tidigare ansågs värdefull blivit dussinvara. Området informationshantering är däremot fortfarande i snabb omdaning, dels med aktörer som tillkommer, läggs ner eller köps upp, dels är under inflytande av en snabb teknikutveckling.

På konferensen rörde de allra flesta inslag informationshantering (webbdesign, söksystem, informationsarkitektur etc) medan onlinetjänster, databaser, e-publikationer, online-/informationsbranschen som sådan, etc spelade en mycket undanskymd, för att inte säga nästan obefintlig, roll. (Det här tycks vara ett generellt fenomen, i alla fall att döma av referat från andra konferenser i år, bl a SLAs i somras, där "det nästan bara talas om teknik och 'solutions' men inget om 'content' ").

Vad gäller utställningen framhävde arrangören just informationshantering, bland annat genom att inte bara upprepa förra årets nyhet, dvs "utställningen i utställningen" CME (Content Management Europe), utan också lansera ännu en sådan "utställning", nämligen EDRM (Enterprise Document & Records Management Europe). Faktum är att nästan lika många utställare var inriktade på informationshantering som på information. Och sammanställer man utställarnas egen kategorisering (de fick välja *en eller flera* bland 13) blir bilden som nedan. Även om man kan undra över somliga utställares val av kategori/er säger den i alla fall en del.

Kategori	Antal utställare som uppgav kategorin
Content Management	82
STM	80
Intranets & Portals	64
ePublishing	60
Business & financial	46
Market research & CI	43
Collab. solutions & KM	38
Legal & regulatory	35
Library Manag. solutions	34
Enterprise search	29
Document management	24
Records management	23
Recruitment & training	14



Det må vara hur som helst med detta - ytmässigt var det inget snack om att det var informationsaktörer som dominerade tungt, med sådär tre fjärdelar av ytan. Allra störst monteryta hade giganter som Thomson och Wolters Kluwer, onlineaktörer som STN/FIZ Karlsruhe/CAS, förlag som Springer och Wiley och slutligen EU-kommissionen. Elsevier då? Nej, de hade mindre yta (men märktes lika bra ändå, mycket beroende på uppmärksamheten kring dess nyligen lanserade jättedatabas Scopus). Apropos Elsevier: LexisNexis uteblev även i år, för andra året i rad.

## Googlemania

Google deltog inte som utställare och hade inte heller någon föredragshållare på konferensen. Ändå kändes Google högst närvarande. Precis som i verkliga livet vad gäller onlinebranschen vilade nämligen Googles ande över Online 2004. Det gällde konferensen, där Google fördes på tal i *samtliga* föredrag jag bevistade, oberoende av inriktning. Och det gällde utställningen, där Googles påverkan var högst påtaglig.

I det senare fallet syftar jag på den trend hos de stora aktörerna inom online vad gäller att ge samlad tillgång till allt större delar av sitt utbud av information - och att all denna information ska vara samordnat och enkelt sökbar via en "Google-sökruta". Se bara på exempel hos de största aktörerna: Elsevier - Scopus. CSA - Illumina (dess nya gränssnitt med sökning i allt material). Thomson Gale - på väg mot gemensam plattform, med samordnad sökning. Wolters Kluwer - på väg mot en "Single Internet Platform" vad gäller den i dag spridda rättsinformationen. M fl. Så även i detta avseende kan man säga att Online 2004 var en god spegling av verkligheten.

Det ska sägas att bakom Googles "närvaro" låg inte bara det "normala" intresset för Google, med betoning på dess webbsökmotor (och förstås Google Scholar, den nya sökmotorn för sökning i akademiskt och vetenskapligt material), utan också företagets nylansering av sin enterprise-sökmotor, Google Search Appliance, i kraftigt förbättrad version. Denna hårdlanseras nu i Europa, och som ett led i detta bevisades Online 2004 av Dave Girouard, chef för Googles satsning på enterprise-sökmotor. Trots att Google alltså inte ställde ut gavs han möjlighet att hålla en produktpresentation på utställningen - samt att disponera ett rum i anslutning till presscentret för intervjuer. Så därför var Google faktiskt också lite "fysiskt" närvarande! Försedd med presspass tog jag f ö tillfället i akt att genomföra en intervju med Dave Girouard. Den följer i avsnitt 5.

## ***Onlines roll***

Den som läst tidigare års rapporter från Online har kunna notera mina resonemang kring evenemangets roll. I rapporten från Online 2003 skrev jag att "Online Information inte längre är branschens centralpunkt (... och att detta är helt naturligt eftersom ...) evenemanget helt enkelt speglar onlinebranschens utveckling. 1977 var det frågan om en synnerligen ung bransch: de första allmänt tillgängliga onlinetjänsterna hade ju startat bara 5 år tidigare. Under de följande åren etablerade sig Online Information som en internationell centralpunkt för branschen, där det var självklart att såväl de viktigaste aktörerna som personerna deltog. Nu 27 år senare är branschen inte bara många, många gånger större utan det rör sig också om en "mogen" bransch som inte tycks behöva denna typ av centralpunkt."



Men nu efter ytterligare ett år så kan jag notera att Online ändå *är* så nära centralpunkt som det går i en bransch som trots allt omsätter ett par hundra miljarder kr. Det säger en del att av de 100 företag som Information Today anser "matter most in the digital content industry", enligt sammanställningen i [www.econtentmag.com/EContent100](http://www.econtentmag.com/EContent100), så ställde minst 27 ut på Online 2004, nämligen Alacra, anacubis, Arbortext, British Library, CAS, CSA, EBSCO, Ektron, Elsevier, EMC (Documentum), eMeta, Endeca, Euromonitor, Factiva, FAST, HighWire Press, D&B (Hoover's), Ingenta, IXIASOFT, Knovel, Ovid, Percussion, ProQuest, Questel Orbit, RedDot, Reuters (Business Insight), Springer, Swets och Thomson.

Orsaken kanske är att det inte finns någon annan konferens med såväl fokus som bredd inom området. Information Today lade ju i år ner sitt National Online Meeting (USA) efter 23 år, och de evenemang som man ersatt detta med, liksom andra arrangörers evenemang, har inte alls samma omfattning och bredd som Online, varken vad gäller innehåll eller deltagare. Samtidigt har den senaste tidens återhämtning inom de delar av onlinemarknaden som på senare år haft det lite svårt gynnat "övervintraren" Online.

Intressant nog ägnade Information Today stor uppmärksamhet åt Online 2004, med inte bara monter utan också med en blogg ([www.infodayblog.com](http://www.infodayblog.com); värd att ta en titt på - mycket bilder! Hur länge

innehållet finns kvar vet jag dock inte), som löpande fylldes på med rapporter, kommentarer och foton från utställningen. De producerades bl a av närvarande onlinekändisar som Dick Kaser, Marydee Ojala, Nancy Garman och Paula Hane.

Här kan f ö nämnas att den delvis konkurrerande mässan KM Europe, som ägt rum i Amsterdam i november de senaste 5 åren (och som nog delvis är orsaken till att VNU så hårt betonar informationshantering när man marknadsför Online) i år inte tycks ha varit särskilt lyckosam, i alla fall enligt vad jag hört. KM Europe 2004 hade ca 50 utställare, dvs färre än förra året, och besöktes av 2-3.000 personer. Till skillnad mot Online är konferensdelen på KM Europe gratis - men å andra sidan mer produkt- och PR-inriktad.

För att återgå till Online så vill jag återigen påminna om att evenemanget inte längre (som "förr i tiden") är "*det*" tillfälle aktörerna väljer för sina viktigaste lanseringar. De relativt sett ganska få pressreleaser etc som presenteras är ofta av tidigare datum eller (ny)lanserar redan befintliga tjänster, funktioner, etc.

Med detta sagt vill jag ändå säga som jag brukar: Om man bara har möjlighet att bevista en konferens/utställning per år så är det, enligt min åsikt, Online man ska välja.

### ***Online 1999-2004 - lite statistik***

Det är nu sjätte gången jag rapporterar från Online i temanummer av Online- och webbnyheter. Innan jag går vidare med detaljer från konferens och utställning vill jag passa på att presentera en liten sammanställning av uppgifter från mina rapporter i Online- och webbnyheter under dessa år, inkl antal delegater från de länder som nyhetsbrevet distribueras till (dvs Sverige och Finland). Allt med reservationer för eventuella missar ...

	<b>Konferensen</b>					<b>Utställningen</b>		
	<b>Delegater</b>			<b>Föredrag *)</b>		<b>Utställare</b>	<b>Besökare</b>	
	<i>Totalt</i>	<i>Sverige</i>	<i>Finland</i>	<i>Totalt</i>	<i>Sverige</i>	<i>Totalt</i>	<i>Sverige **)</i>	<i>Totalt</i>
<b>Online 1999</b>	<b>1.000</b>	<b>&lt; 60</b>	<b>&gt; 60</b>	<b>60</b>	<b>1</b>	<b>300</b>	<b>1</b>	<b>&gt; 20.000 ***)</b>
<b>Online 2000</b>	<b>1.000</b>	<b>58</b>	<b>75</b>	<b>75</b>	<b>0</b>	<b>300</b>	<b>1</b>	<b>&lt; 20.000 ****)</b>
<b>Online 2001</b>	<b>750</b>	<b>25</b>	<b>50</b>	<b>80</b>	<b>1</b>	<b>290</b>	<b>1</b>	<b>12.700</b>
<b>Online 2000</b>	<b>&gt; 700</b>	<b>34</b>	<b>33</b>	<b>60</b>	<b>1</b>	<b>240</b>	<b>2</b>	<b>11.000</b>
<b>Online 2003</b>	<b>&lt; 700</b>	<b>27</b>	<b>33</b>	<b>60</b>	<b>1</b>	<b>250</b>	<b>1</b>	<b>11.000</b>
<b>Online 2004</b>	<b>650</b>	<b>27</b>	<b>36</b>	<b>90</b>	<b>1</b>	<b>250</b>	<b>3</b>	<b>10.600</b>

< Något under

> Något mer än

\*) Avser föredrag/papers (ej paneldiskussioner m fl "inslag") på konferensen (betaldelen)

\*\*) Ej medräknat svenskt deltagande i icke-svenska montrar

\*\*\*)) Störst antal någonsin sedan starten 1977

\*\*\*\*)) Uppskattning

## **2. Konferensen**

### ***Delegater: Finland på tredje plats***

Antalet betalande deltagare på konferensen var alltså ca 650, dvs ytterligare något färre än vid Online 2003, då strax under 700 deltog. Ca 40 länder var representerade. Flest delegater kom som vanligt från UK, 266 st.

Därefter kom USA med 42, Finland 36, Tyskland 34, Norge 32, Nederländerna 30, Sverige 27 (varav 14 från Informationsspecialisterna, dvs ungefär som förra året), Frankrike 19, Schweiz 18, Danmark 14, Belgien 13, Ungern 12, Irland 9, Spanien 7, Portugal 7, Italien 7, Canada 7, Tjeckien 6, Österrike 6, Japan 5, Luxemburg 5, Slovenien 5, Sydafrika 5, Chile 4, Israel 4, Zimbabwe 4, Australien 3, Island 3, Singapore 3, Indien 2, Iran 2, Grekland 2, Polen 2, Egypten 1, Nigeria 1, Sydkorea 1, Kuwait 1 och Ukraina 1. Vad gäller såväl Sverige som Finland kom närmare två av tre från offentlig sektor (myndigheter, utbildning och andra offentligt stödda organ).

### ***Webbdesigngurun Jakob Nielsen - en publikmagnet***

Somliga deltagare väljer att bara vara med (och betala för) för en dag. Att de allra flesta av dessa valt konferensens första dag rådde det ingen tvekan om. Denna inleddes nämligen med en fullsatt, eller snarare överfull, salong. Orsaken var att konferensens verkliga publikmagnet, webbens "king of usability"/webbdesigngurun Jakob Nielsen, då höll sin presentation "What's new, what's old in online usability". Även i övrigt var denna dags program vara mest attraktivt, i alla fall enligt min åsikt och enligt andra jag träffade.

Just "usability" var för övrigt den ansats som **Martin White**, Intranet Focus (UK), konferensens ordförande för femte gången (... och deltagare för 28:e gången!), hade när han öppnade konferensen. Han konstaterade det "märkliga i att det trots århundradens erfarenhet av 'usability' (här visade han en bild på en gammal kyrkorgels avancerade men användarvänliga (?) klaviatur) fortfarande förekommer så många dåligt utformade 'användargränssnitt' - för allt från webbplatser till hemmets fjärrkontroller".

Martins ansats passade bra inte bara som en introduktion till Jakob Nielsens föredrag utan till hela konferensen. Den visade sig nämligen ha en mycket stark slagsida mot sådant som har med "användning" (i bred bemärkelse, inkl informationshantering) att göra, i motsats då till "information" och "inhåll". I storleksordningen 80 % av sessioner och föredrag behandlade nämligen KM, CM, IA och "enterprise search" (sökmotorer/söksystem) samt, i mindre utsträckning, usability/gränssnitt, intranät, portaler och övrig "praktik". Den resterande femtedelen hade teman som affärsinformation, vetenskaplig publicering/OA, sökning på nätet och lite annat som åtminstone med lite god vilja kan sägas tillhöra information/innehåll. Föredrag som behandlade "vanliga" onlinetjänster och databaser förekom i princip inte alls. Faktum är att information/innehåll i år hade en *ännu* mer underordnad roll än vid förra årets Online.

Konferensens uppläggning var sig lik. Den löpte i tre parallella sessioner, normalt med halvtimmesföredrag men med några längre. Någon enstaka paneldiskussion förekom också - dock färre än brukligt och, enligt min mening, alltför få. Det totala antalet föredrag var ca 90 st - vilket är väsentligt fler än vid såväl Online 2002 och Online 2003 då antalet var ca 60. Orsaken var att en del av tiden då upptogs av 5 st en och en halvtimmes "minikurser", benämnda "Online Information Academy", vilket man nu lämnat (rättare sagt ersatt med entimmes s k "Roundtables"-seminarier under lunchpausen).

### ***Tre svenska föredrag!***

Kvaliteten på föredragen vad gäller framförandet var genomgående god, för att inte säga mycket god. Priset för detta var tyvärr i likhet med de närmast föregående åren en mycket stor geografisk snedfördelning: en mycket stor del av föredragshållarna kom från UK och USA. Men för en gångs skull var Sverige rejält representerat, med så många som tre föredrag! Det kan jag inte minnas förekommit någon gång tidigare. De tre föredragshållarna/föredragen var *Dick Stenmark* (Informatik, Göteborgs universitet): "Managing knowledge through everyday activities", *Mikael Thorson* (InfoSphere AB): "Choosing the perfect search engine for your organisations" och *Håkan Axelsson* (EuroInfoPool/-InfoData): "Official business information and the expanding EU challenges & opportunities". Referat följer i avsnitt 6.

Vad gäller föredragens innehåll upplevde jag att det, generellt sett, i ganska stor utsträckning rörde sig om presentationer på nivån "kurs" eller "utbildning" snarare än resultatet av "forskning" eller i alla fall "nytt" inom området. Förutom konkreta exempel, som t ex föredrag om hur man väljer sökmotor för sin organisation, om vad RSS är för något och om kvalitetskrav på information på webben, har jag har lite svårt att sätta fingret på detta utan kan bara konstatera att nutidens krav på föredrag och papers uppenbarligen är ganska så annorlunda än de som tillämpades åtminstone fram till mitten av 1990-talet.

Jag var själv granskare av papers inför Online 1993. För att få ett paper antaget ställdes på den tiden bl a kravet att dess innehåll inte skulle ha publicerats tidigare. Detta gäller uppenbarligen inte längre. Kravet på att föredragen inte får utgöra produktreklam tycks dock ha bibehållits; bara ett fåtal av föredragen vid Online 2004 kunde betraktas som sådana. Det tackar vi för!

### ***40 av 90 föredrag inte dokumenterade***

Vad beträffar konferensdokumentationen hade arrangören rationaliserat sitt eget arbete rejält genom att dela ut proceedings i tryckt form vid registreringen - och "that's it"! Visserligen lades dokumentationen efter konferensen ut på webben - men den bestod av exakt de papers (i pdf) som finns i den tryckta volymen, dvs inga kompletteringar med t ex visade powerpoint-sviter eller av saknade papers. Precis som "förr i tiden" alltså - efter ett antal år med allt från dokumentation på CD-ROM till webbversioner som blivit tillgängliga i efterhand, ibland lösenordskyddade. Pdf-versionerna av årets papers, som alltså inte är lösenordskyddade, finns tillgängliga via [www.online-information.co.uk/2004proceedings](http://www.online-information.co.uk/2004proceedings). Trots att detta alltså är gratis kostar tryckta proceedings för den som inte deltog i konferens så mycket som 1.500 kr!

Tyvärr finns bara ca 50 av de 90 föredragen dokumenterade i form av papers (i proceedings och på webben; många är dock väldigt kortfattade). Med tanke på den ganska digra konferensavgiften, som lägst ("Super Early Bird") ca 9.000 kr, anser jag att man faktiskt kan begära mer - i synnerhet som publicering av papers i efterhand från "försumliga" föredragshållare ju är hyfsat enkel att åstadkomma tack vare webben.

### ***Svensk nominering på Awards Dinner***

Konferensavgiften innefattade som vanligt konferensmiddagen, vilken ägde rum andra konferensdagens kväll. Den benämndes "2004 International Information Industry Awards" - ett bra namn med tanke på temat, dvs utdelande av priser - *många* priser: totalt 15 st, sponsrade av diverse företag. Några exempel: Innovation in CM, Best Intranet or Extranet Project, Best STM Information Product, Best User Experience, IWR Information Professional of the Year och Lifetime Achievement Award.

För ovanlighetens skull fanns en svensk nominerad, nämligen i klassen Best Business Information Product. Det var InfoDats *EuroInfoPool*, som f ö var ende svenska utställaren på Online i år. EuroInfoPool tillhandahåller en portal för samlad åtkomst till offentliga bolagsregister i Europas länder samt kreditinformation om 25 miljoner företag. Men det blev konkurrenten Hoover's Online som knep förstaplatsen.

Apropå att Googles ande vilade över Online kan här tilläggas att i klassen Best Search Product så delades priset mellan Fast ESP, från norska sökmotorföretaget FAST, och Google - som dock inte hade någon representant på plats för att hämta priset.



### 3. Utställningen

Som nämnts var antalet utställare ca 250, varav närmare hälften inom "informationshantering". Dessa senare upptog dock bara en mindre del av utställningsytan eftersom flertalet bara hade små montrar. Kanske tre fjärdedelar av ytan upptogs av informationsaktörer, trots att de till antalet var kanske hälften. Det här säger en del om aktörernas storlek eftersom monterpriserna är avsevärda. En normalstor monter om ca 12 kvm betingade ett pris om sådär 50.000 kr - ganska häftigt för många av aktörerna inom informationshantering. Aktörer med de största montrarna, bland dem Thomson, Wolters Kluwer och EU-kommissionen, gissar jag däremot ha fått spendera i storleksordningen uppemot någon miljon kr var, inkl kostnad för presentationer, för sitt deltagande.

Utöver ovan nämnda kategorier utställare deltog organisationer som CILIP, Society of Archivists och Association of Independent Information Professionals (AIIP), EUSIDIC och SLA samt "branschpublicister" m fl som Information Today och FreePint.



Jag har ingen uppgift om hur många av utställarna som tillkom resp. försvann jämfört Online 2003, men jag gissar på ett 50-tal. Av sådana som försvunnit noterade jag i synnerhet H.W. Wilson/Wilson-line och FT Corporate Solutions/FT.com. I och med att LexisNexis slutat vara med på Online har också rättsinformation i princip helt försvunnit - med det stora undantaget Wolters Kluwer. Även Thomson har dragit ner på sin rättsinformation. Bland andra jag skrev om förra året saknades nu Global Securities Information, Mergent, Standard & Poors, Productscan Online, thebigword, BBC Monitoring, RDN och så förstås Micropatent (uppköpt av Thomson) och Nextpage (uppköpt av FAST).

Närvaron av aktörer utanför USA och UK blir gradvis allt tunnare. Bara Frankrike och Tyskland deltar väl i samma omfattning som tidigare - Tyskland pga färre men större aktörer: FIZ Karlsruhe/ FIZ Chemie/STN International och Springer. Aktörer från länder utanför Nordamerika och Västeuropa är väldigt få. Men det är klart - räknar man de "internationella" aktörerna som nationella (t ex Reed-Elsevier som holländskt och Thomson som kanadensiskt) så blir ju bilden helt annorlunda ....

Kreditinformations-/företagsinformationsaktörer typ ICC och D&B har lämnat Online redan tidigare, liksom de med finansiella tjänster (t ex Bloomberg). De "rena" söksystemsaktörer som var med för några år sedan (t ex Autonomy, Verity, ZyLab) har lämnat scenen för FAST, Endeca och Mondosoft. Aktörer med "överliggande" system (typ visualisering) ser man inte mycket av numera. E-böcker och bibliotekssystem var litet nu som tidigare. Vad har då tillkommit? Ja, det är väl lösningar för webb, intranät, publicering mm inom "informationshantering".

#### ***Inte bara montrar***

I anslutning till utställningen pågick *väldigt* mycket annat. Där hölls t ex ett stort antal, drygt 110 (!), gratispresentationer, normalt halvtimmeslånga. OK, en del av dem var dubblerade, men ändå ... De var grupperade i seminarier inom områden som "Content", "Content management", "Enterprise document and records management", "Technology case studies" och "E-publishing solutions", och var ofta men långt ifrån alltid produktpresentationer. Till exempel bestod den sista dagens seminarium inom "information skills and career development" av presentationerna "Core competences for information professionals", "Alternative career paths for information professionals", "Top 10 tips for information professionals working in records management", "10 techniques for advanced web searching", "How to

sell your brand" och "RSS for beginners". Och så "Enterprise search solutions" där Googles Dave Girouard höll sin dragnings.

Och så förekom s k "Academy sessions", en och en halvtimmes seminarier ledda av TFLP, vilka förmedlade kunskaper inom områden som "Communicating your purpose" och "Creative conversations". Det avhölls också ett antal sk Association Events (vissa i kombination med luncher), anordnade av sammanslutningar m fl som ALPSP, SCONUL, CILIP/ICLG, EUSIDIC/NFAIS, Aslib/KIMNET, AIIP, SLA och Business Statistics User's Group.

### ***Nordic Forum***

Det anordnades också s k International Forums, en timme och en kvart långa, bland dem Nordic Forum. Jag hade inte möjlighet att bevista det, men enligt programmet hölls här presentationer av *Gunilla Widén-Wulff* (Åbo Akademi University), *Ulla Nygren* (Åbo Handelshögskola) och *Christina Tovoté* (Stockholms univ.bibl.). Christina talade om det nordiska nätverket för informationkompetens. Forum anordnades också för Frankrike, Spanien, Nederländerna, Tyskland och Italien.

Ja, inte behöver man vara delegat på konferensen för att ha full sysselsättning under - och säkert stort utbyte av - Online! Jag vet inte hur många svenskar som väljer just detta sätt, men jag gissar på uppemot ett hundratal, vilket skulle innebära ca fyra gånger så många som antalet konferensdelegater.

## **4. Utställare**

250 utställares montrar går såklart inte att besöka på tre dagar, i synnerhet om man som jag också har konferensen att tänka på. Precis som tidigare års Online koncentrerade jag mig därför på utställare med "information". Däremot gjorde jag i år avkall på tidigare års ambition att täcka in så mycket som möjligt av intresse inom denna kategori. Jag valde istället att primärt besöka stora aktörer med onlinetjänster men i övrigt bara göra lite småplock.

Jag tror ändå inte att jag missat så många rena nyheter, i alla fall att döma av de relativt fåtaliga pressreleaser som jag samlade på mig. Och de intressanta nyheter jag kan ha missat bör dyka upp i ordinarie Online- och webbnyheter. Vad gäller beskrivningar om utställarna går det ju f ö bra att anlita den kompletta utställarförteckningen på [www.online-information.co.uk/cgi-events/exhibitors.pl?exhibition\\_id=66](http://www.online-information.co.uk/cgi-events/exhibitors.pl?exhibition_id=66). Där finns också länkar till resp. utställares webbplats.

### ***Thomson***

I *Thomsons* stora montergrupp besökte jag Thomson Gale och Thomson Dialog. Det jag intresserade mig mest för hos *Thomson Gale* var den kommande "plattformen" för all dess information: Thomson Gale Powersearch. Hittills nyttjas den för InfoTrac-produkterna. Samtliga övriga produkter kommer att migreras dit, med början år 2005. År 2006 ska allt vara klart, och då ska man kunna utföra sökning i allt Gales material, totalt omfattande 8 TB, via ett enda gränssnitt. ("Googlemania" ...). Sökresultat från de olika delarna i Gales utbud kommer dock alltid att presenteras under separata "flikar" - eftersom det rör sig om så olika typer av material: från referensverk till artiklar, från nyheter till multimedia. Man riktar sig i första hand mot den





akademiska sektorn och bibliotek.

Vad gäller **Thomson Dialog** tittade jag främst på den nya tjänsten Dialog Live News, vilken jag beskrev i förra nyhetsbrevet (2004:6). Det är alltså en tjänst som i realtid levererar nyheter inom användarens eget valt område (branscher, företag etc) till den egna PC:n. Den baseras på ett nyhetsflöde (främst affärsinformation: marknader, branscher, företag, politiska händelser, ...) i realtid från 2.300 källor direkt från informationsleverantörer (nyhetsbyråer, tidningar, tidskrifter etc). Live News är en separat tjänst som också kan nyttjas som tillägg till befintligt abonnemang på Dialog NewsRoom eller Dialog Profound.

Jag passade förstås på att fråga hur det går med migreringen av tjänster till den nya plattformen. I våras togs ju de två första tjänsterna i drift där, nämligen Dialog NewsRoom och Dialog Profound. Då sades det att de övriga tre som skulle ingå, Dialog Intelliscope, Dialog Company Profiles och Dialog NewsEdge, skulle följa senare under 2004. Så har inte blivit fallet, och den jag pratade med i montern sade att Company Profiles och NewsEdge skulle migreras "om 4-5 månader", dvs i april-maj 2005. Samtidigt läser jag i Dialog Chronolog att samtliga kunder till Company Profiles i oktober flyttades över till Thomson Gale Business and Company Resource Center (BCRC), som innehåller 440.000 företagsprofiler om företag runt om i världen. (Jag kollar upp hur det egentligen är och återkommer i senare nyhetsbrev).

DataStar och "gamla" Dialog då? Det har ju från tid till annan spekulerats i en kommande sammanslagning av dessa till en tjänst. Svaret jag fick var att det för närvarande inte är aktuellt, men inte uteslutet för framtiden. Det är för övrigt intressant att notera att DataStar börjat lyftas fram lite mer än på länge. Det är åtminstone en känsla jag fått. Och då menar jag såväl i marknadsföringen (där tjänsten förs fram som den rätta källan för informationssökning inom vetenskap och teknik) som vad gäller funktionalitet (nu med nya bevakningsfunktioner, nya möjligheter för efterbearbetning av sökresultat, mm).

Thomson slog på stora trumman för sin helt nya produkt **Thomson Pharma**, som lanserades just på Online 2004. Jag hann tyvärr inte studera den så här bara lite info från pressmaterial mm. Thomson Pharma ger användarna möjlighet att följa ett läkemedel under hela utvecklingsprocessen. Materialet hämtas från andra produkter och tjänster inom Thomson (grunddata, nyheter, patent, artiklar, företagsfakta, mm). Thomson Pharma, som blir tillgänglig i januari 2005, ingår i **Thomson Scientific**, från vars övriga sortiment också bl a visades ISI Web of Knowledge, Web of Science, Delphion, Derwent m fl. Betr Web of Science annonserades att projektet "Century of Science" snart skulle ge konkret resultat i form av en komplettering av referenser från åren 1900-1944. Även ISI Web of Knowledge kommer att få del av detta. Apropå patent har Thomsons köp av IHI, med bl a konkurrenten MicroPatent, nu godkänts av myndigheterna.

Något Thomson inte hade med i år var Thomson Legal & Regulatory Europe, där ju bl a Westlaw ingår. Den viktigaste orsaken bör förstås vara att konkurrenten LexisNexis inte skulle delta (vilket Thomson säkert visste på förhand, eller i alla fall antagit eftersom de inte var med förra året). Å andra sidan var Thomson rättsinformation ganska nedtonad redan förra året. Och nu tycks det som om Online inte längre alls attraherar aktörerna inom detta område. Utöver Wolters Kluwer deltog, såvitt jag såg, inga andra. För bara några år sedan var det ett ganska "hett" område på Online.

Slutligen kan betr Thomson berättas att bland de fem "International InfoStars" man utsåg för 2004 fanns finländska *Minna Abrahamsson*, Nokia.

## **Elsevier**

Som jag antytt var **Elseviers** huvudnummer den nyligen lanserade jättedatabasen **Scopus**. Och uppmärksamheten kring denna var också mycket stor, med ofta mycket folk i montern. Annat från Elsevier (t ex Scirus, MDL och Emerald) kom helt i skymundan för detta - vilket säkert var såväl

önskat som väntat. Jag skrev om Scopus i det förra nyhetsbrevet (2004:6). Men eftersom det blir lite "orättvist" mot Elsevier (dvs jämfört med Thomson ...) om jag bara slutar här så upprepar jag lite.

Scopus innehåller dels 27 miljoner referenser till artiklar mm i 14.000 vetenskapliga publikationer från totalt 4.000 förlag, dels 167 miljoner webbsidor med vetenskaplig information (från Scirus). Vissa publikationer täcks ända fr o m 1966. Referenserna består av bibliografiska uppgifter och i förekommande fall abstracts. De hämtas bland annat från Science Direct och databaser som Embase, Compendex, PsycInfo, Medline m fl. Referenser fr o m 1996 är också försedda med citerade referenser. Referenserna innehåller också länkar till fulltext, vilken kan nås i den mån användarens bibliotek/organisation abonnerar på publikationen ifråga.

Jag kan väl tillägga att den demo jag fick visade att även om man kan söka i allt material samtidigt ("Googlemania" ...) så fördelas sökresultaten alltid på olika discipliner (presenteras i flikar). Varför? Jo, vid de användarstudier man gjort visade det sig att det blev alltför förvirrande att presentera allt i en enda lista à la Google. Hur skulle materialet sorteras i så fall? Men utvecklarna hade haft många diskussioner om detta innan de slutligen bestämde sig!

## ***Wolters Kluwer: Rättsinformation***

**Wolters Kluwer** hade som nämnts en av utställningens allra största montrar/montergrupper. För några år sedan uttalade Wolters Kluwers ledning bekymmer över att deras företag låg på efterkälken vad gäller online/Internet, jämfört med t ex Thomson och Elsevier. Utöver Ovid, som man köpt ett par år tidigare, fanns bara spridda "online/Internet-öar" inom koncernen. Man bestämde sig för att framdeles intensifiera satsningen på Internetrelaterade produkter. Jag upplever nu att man börjat få upp farten. Man ska väl inte dra för stora växlar på närvaron på Online, men i detta fall upplever jag i alla fall att den är en spegling av "verkligheten". Det har börjat hända saker, bl a vad gäller rättsinformation.



Vad gäller rättsinformation är Wolters Kluwer en av världens stora. Ändå har man inte lyckats samla eller samordna denna i stil med vad Thomson (Westlaw inkl Sweet & Maxwell m fl) resp Elsevier (LexisNexis inkl Butterworths m fl) gjort. Speciellt stora är man i Europa, med en omsättning år 2003 på ca 12.5 miljarder kr. Som man själva skriver: "Wolters Kluwer Legal, Tax & Regulatory Europe's market positions are significant, and it is the number one or two player in almost every core market and geography". Men ändå har dess förlag som Norstedts Juridik, Kluwer (Nederländerna, Belgien), Lamy (Frankrike) etc agerat själva - och ganska långsamt - vad gäller online (med separata tjänster som Colex Data, jura.be, Lamyline Reflex, Kluwer, Croner-i, chhinformation.com, SRL, m fl). En samordnad satsning benämnd Juriforum, annonserad kring 2001, kom av sig.

Men nu tycks en ordentlig uppryckning vara på gång. En gemensam teknisk plattform, Wolters Kluwer Internet Platform (EIP), är på väg att tas fram. Den ska utgöra en plattform för publicering såväl som för "portaler", dvs onlinetjänster. Gemensam sökning i alla databaser och tjänster ska medges ("Googlemania" ...). Ambitionen är helt enkelt att denna "Single Internet Platform" ska göra Wolters Kluwer till en "global player"! Man har inlett med Nederländerna, Belgien och Spanien, fortsätter under 2005 med Italien, Tyskland, UK och USA och slutligen ("efter 2005") med Sverige, Norge, Danmark och Centraleuropa.

Jag frågade "givetvis" i montern hur man tänkt sig det här med söksystem. Många Kluwer-företag inom rättsområdet (bl a CHH och Norstedts Juridik) nyttjar idag NXT-plattformen från Nextpage - som ju i höstas köptes av FAST. Den person jag slutligen fick prata med - och som jag uppfattar som

auktoritativ - berättade (i min tolkning!) att man börjat arbeta med NXT men nu efter FASTs köp av Nextpage samarbetar med FAST i syfte att ta fram en lösning som har FAST i botten men där gränssnitt som är mer anpassade till NXT (vars gränssnitt ju i mycket lämpar sig just till rättsinformation) ska kunna nyttjas. Denna information passar ihop med den jag fick i samtal med vd för FAST och fd vd för Nextpage vid mitt besök i FASTs monter. (Mer om detta nedan under FAST).

### ***Wolters Kluwer: Ovid (inkl Silverplatter)***

**Ovid** fortsätter att utvecklas snabbt. Om det vittnar inte minst att antalet databaser online vuxit till ca 300 (inkl de som finns vid **SilverPlatter**). Bland det man talade mest om i Ovids monter var förstås den senaste tidens nyheter kring Books@Ovid, som nyligen gått över till en ny teknisk plattform och därtill i januari 2005 utvidgas med Martindale (som tycks bli en separat produkt). Men det här har jag ju redan skrivit om i tidigare nyhetsbrev (2004:6) så jag förhörde mig istället (som vanligt) om planerna för SilverPlatter.

SilverPlatter tycktes ju efter köpet 2001 vara på väg att integreras i Ovid även "tekniskt". Så här skrev jag dock i nyhetsbrevet 2004:6: "Eventuella antaganden om att SilverPlatter skulle upphöra som egen produkt och att dess innehåll skulle integreras i Ovid tycks komma på skam. Ovid berättar att man under det senaste året satsat 70 miljoner kr på SilverPlatter, vilket bl a lett till att International Science Database (ISD) lanserats."

Den representant för Ovid jag nu pratade med "erkände" nu också att man för "2-3 år sedan verkligen hade tanken på att lägga ner SilverPlatter och integrera dess innehåll med Ovid". Detta är också orsaken till att varumärket rensades bort. Men så för 1-2 år ändrade man sig: SilverPlatter ska bestå åtminstone de 5-6 närmaste åren. Hur det blir därefter får man se. Blir det då någon "ämnesmässig" uppdelning, med Ovid = Health och SilverPlatter = Humanities? Nej, någon sådan inriktning har Ovid inte.

### ***FIZ Karlsruhe/STN/CAS/FIZ Chemie***

**FIZ Karlsruhe/STN/CAS/FIZ Chemie** firade STNs 20-årsjubileum med tårta under Onlines första dag. Även Derwent, som ju har databaser vid STN, deltog på ett hörn i firandet - som samlade mängder med folk. FIZ Karlsruhes chef Sabine Brünger-Weilandt höll tal, och hon gjorde bara en förutsägelse: Om ytterligare 20 år, 2024, ska STN fira sitt 40-årsjubileum "och då är ni alla välkomna hit igen för att fira det tillsammans med oss!"

Men det var förstås inte allt, utan som vanligt på Online presenterades mycket nytt - nåja, det mesta var inte helt nytt, däremot var det *mycket*. I de två senaste Online- och webbnyheter (2004:6 och



*Trångt när det firades ...*

2004:5) finns notiser om sju av nyheterna: "Fortsatt uppåt för FIZ Karlsruhe och STN" (om den positiva årsrapporten), "FIZ Karlsruhe och Max Planck Society får statligt stöd för att bygga plattform för vetenskaplig kommunikation", "Ännu fler CSA-databaser nu online vid STN", "Drug Data Report (PROUSDDR) ny läkemedelsdatabas vid STN", "Ännu fler CSA-databaser nu online vid STN", "CA tillförs ytterligare äldre material" (om utvidgningen av projektet Science Century) och "FIZ Karlsruhe och FIZ Chemie sammanför sina arkivaktiviteter".



*... men tårtan var god!*

I övrigt presenterades bland annat "STN Express with Discover!", som bistår med analyser och visualisering (inkl diagram i 3D) av sökresultat, samt FIZ AutoDoc, som erbjuder möjlighet att beställa dokument från "de ca 450 miljoner vetenskapliga publikationer som finns i form av referenser i STNs 220 databaser". FIZ Chemie berättade också bl a om sin nya databas POLYRESIST.

## CSA

Att CSAs, *Cambridge Scientific Abstracts*, huvudnummer på Online skulle vara CSA Illumina var inte så svårt att gissa - med tanke på all förhandsinfo. Jag nämnde den också kort i nyhetsbrev 2004:6. Vad är då CSA Illumina? Jo, det är en helt ny tjänst för CSAs 75 databaser (egna och licensierade). På Online förhandsvisades den, men precis när detta skrivs i mitten av december släpps den till kunderna. All information från alla databaser kan sökas via en enda sökruta ("Googlemania" ...). Sökresultat redovisas under en flik benämnd "All Publication Types" (där det också finns uppgift om antal träffar) men bredvid finns flikar med sökresultatet uppdelat på typer av publikationer, t ex "Journals", "Peer-Reviewed Journals", "Conferences" och "Web Sites", var och en med uppgift om antal träffar.

Roddy MacLeod ger i Internet Resources Newsletter ett mycket gott betyg till CSA Illumina: "tänk om andra söktjänster var lika enkla att använda!". Användarna av CSAs befintliga CSA Internet Database Service (IDS) får nu tiden fram till den sista januari 2005 på sig att flytta över till Illumina.

## Factiva

**Factiva** hade även i år sin monter lokaliserad till CME-delen, vilket bara är logiskt med tanke på att man fortsätter att framhäva sin roll som "integrator" snarare än tillhandahållare av information. Som man själva säger: "Factiva's products and services help companies integrate critical business information into their daily workflow, combining the organisation's internal information with content from Factiva and other sources." Självt tycker jag det är lite märkligt, men kanske beror det på att Factiva hellre än att stängas "öppet" mot sina stora konkurrenter Thomson Dialog och LexisNexis, som ju har mycket mer information, väljer detta något annorlunda spår.



Apropå detta så passade jag på att fråga om det samarbete med just LexisNexis som annonserades nyligen gäller även kunder i Sverige. Svaret var ja. Det innebär alltså att Factivas kunder inom marknadssegmentet advokater/advokatkontor fr o m i mars ska bli kunder till LexisNexis, vars tjänst nu utvidgas med information från de ca 3.400 källor som finns unikt vid Factiva (av dess totalt ca 9.000 källor).

Jag passade också på att fråga hur det gick med det i september förra året annonserade samarbetet med IBM kring dess "plattform" WebFountain, då under utveckling. (WebFountain skulle bli en sorts "analysverktyg" för att finna mönster och trender i mycket stora mängder ostrukturerad information, t ex hämtad från webben. Den skulle Factiva nyttja för en ny tjänst, Reputation Manager, som var tänkt att lanseras under andra kvartalet 2004). Svaret jag fick var att Factiva avbröt samarbetet för ca 3 månader sedan: "Det blev för dyrt; IBM ville ha för mycket pengar av oss".

Apropå nyheter så sade den jag pratade med att en ny tjänst benämnd "Sales Works" (?) ska lanseras i januari. Om jag uppfattade det rätt ska den säljas parallellt med befintliga tjänster (dvs bl a Factiva iWorks). Slutligen frågade jag om läget betr att få DN bland Factivas källor. Svaret var att "DN släpper inte sin information - det är stenhårda på!".



## ***EU-kommissionen samt EPO (European Patent Office)***

**EU-kommissionen** hade tillsammans med **EPO (European Patent Office)** en av de allra största montergrupperna. Här fanns såklart väldigt mycket att ta del av, men jag hann bara besöka den kort. Det jag hann uppfatta av nyheter var följande. EPO kommer att släppa sin nya EPO European Publication Server i januari, med fri tillgång till europeiska patentansökningar och specifikationer (EP-A- resp EP-B-dokument). Den förhandsvisades nu i montern.



Fri tillgång till några tidigare avgiftsbelagda tjänster var också EU-kommissionens huvudnummer. Det gällde EUR-Lex ([europa.eu.int/eur-lex/lex](http://europa.eu.int/eur-lex/lex)), vars nya version (lanserad i november) också tillhandahåller CELEX gratis (vilket förhandsinformerades om redan på Online 2003), samt Eurostat ([europa.eu.int/comm/eurostat](http://europa.eu.int/comm/eurostat)), vars enorma statistikdatabas New Cronos blev gratis i oktober. New Cronos söks via ett nyutvecklat gränssnitt, vilket jag vid min snabbtitt upplevde som mycket användarvänligt.

Bland övrigt nytt var att Europe Direct (informationstjänsten för frågor om hur man hittar dokument på Europaservern, om EUs institutioner mm) som förutom via gratis telefon och e-mail nu även finns tillgänglig via den nya "Web Assistance Service" ([europa.eu.int/eurodirect](http://europa.eu.int/eurodirect)). Den ger möjlighet för frågeställaren att kommunicera "live" via webben med personal vid Europa Direct.

## ***French pavilion, inkl. Questel-Orbit, samt Patanalyst***

Precis som flera gånger tidigare hade ett antal franska utställare samlats i en gemensam montergrupp, benämnd French Pavilion, "The French Electronic Information Industry at Online Information 2004". Ansvarig var GFII, en branschorganisation för producenter, förlag, distributörer, onlinetjänster, förlag



etc, med totalt drygt 90 medlemmar. Bland de 11 som var med fanns vad gäller "information" **BvD Information Management** (konsult och informationsmäklare inom affärsinformation mm), **CCIP** (Paris Handelskammare; erbjuder informationsmäklari vad gäller fransk marknads- och företagsinformation), **INIST** (Institute for Scientific and Technical Information; bl a producent av den multidisciplinära databasen PASCAL) inkl **INIST-CNRS** (French National Center for Scientific Research), **INPI** (Franska patentverket; marknadsför bl a Plutarque, en "smart search services, dedicated to patents, trademarks and industrial design") och naturligtvis Questel-Orbit.

**Questel-Orbit** fortsätter enträget sin satsning på patent och varumärken. Så skrev jag i förra årets rapport från Online, och det stämmer (förstås!) fortfarande. Bland de många patentdatabaserna finns PlusPat, som täcker patent från över 70 länders patentmyndigheter (inkl EPO och WIPO). Den har vuxit till 42 miljoner poster (patent från ansökan till godkännande) och innefattar också bilder, totalt över 10 miljoner st.

Men Questel-Orbit tillhandahåller också "verktyg" av olika slag för att t ex analysera och visualisera sökresultat. På Online passade man på att tillsammans med INPI lansera en ny version av Markush Darc Software för Unix. Den är anpassad för den nya Merged Markush Service, "the largest Markush structure coverage available online with 1.600.000 structures records representing 450.000.000



specific real and prophetic structures". Merged Markush Service produceras gemensamt av INPI och Thomson Scientific.

Ett annat verktyg är PatentExaminer, som nyttjas för avancerade sammanställning av patent. Bland varumärkesdatabaserna finns SWEMARK, svenska varumärken, och FIMARK, finska varumärken. Båda uppdateras veckovis. Utöver patent- och varumärkesdatabaser fortsätter Questel-Orbit att ge tillgång till "veterandatabaser" som CAsSearch, Compendex, INSPEC, Pascal och RAPRA.

En tjänst som jag inte träffat på förut på Online var *Patanalyst*. Bakom tjänsten, som jag uppfattar som en ren konkurrent till Questel-Orbit (och kanske av det skälet *inte* ställde ut i den franska montergruppen), står franska Jouve SA och tyska empolis GmbH. Patanalyst ger tillgång till "all" patent-information: "45 millions of worldwide patents, family databases, abstracts, and full-text documents, as well as facsimile images" (från EPO). Men det man "slåss" med tycks vara de sök-, analys- och visualiseringsverktyg man tillhandahåller (bl a EPOs Viewer). Jouve, med 1.500 anställda, samarbetar sedan länge med EPO (bl a skanningstjänster) och INPI. Priserna för tillgång till Patanalyst börjar på 90.000 kr per år (för 1-4 användare).

### ***Några övriga onlinetjänster inom affärs-, nyhets- och företagsinformation***

*Bureau vad Dijk* upplever jag får en allt starkare position på marknaden för företags- och finansiell information. Man har nu vuxit till 350 anställda och har representanter i 19 länder (där kontoret i Köpenhamn svarar för Norden). Intressant är också att BvD nu till 40 % ägs av investeringsaktören Candover & Cinven, som ju på senare tid köpt på sig allt mer inom informationsindustrin. (Undrar vad som kommer härnäst?).

BvD talade främst om den nya produkten MINT, som lanserades i höstas (tillhandahåller UK-affärs-information), och om helt nya INVIEW, som faktiskt lanserades just på Online 2004. INVIEW tillhandahåller uppgifter om investmentföretag och de börsnoterade företag som de äger investerat i. Den har global täckning med huvudinriktning på Europa och "emerging markets". Precis som andra BvD produkter ingår en rad möjligheter till sammanställning av data, exportfunktioner, verktyg för grafisk redovisning, analyser etc. Bakom INVIEW står förutom BvD företaget Heale Financial, med rötter i bl a Hong Kong och Nya Zeeland och med ett nytt kontor i Argentina. Apropå Norden erbjuder BvD sedan tidigare produkten ODIN med företagsinformation med sitt ursprung bl a i UC, KOB, CreditInform och Suomen Asiakastieto - Finska.

Även *Alacra* tycks stärka sin roll. Den samling databaser som nås via dess olika produkter är imponerande. Bland de ca 100 databaserna finns bland annat Gale Company Profiles, Hemscott Company Reports, Hoover's Company Profiles och Thomson Financial Extel Financials, Moody's Global Credit Research, BvD Zephyr, MergerStat, First Call Estimates, Datamonitor Country Profiles, EIU Country Analysis, Quest Economics, CourtLink by LexisNexis, Edgar Online, Reuters Research, S&P Stock Reports, Investext, Barra Beta Books for Companies, Mintel Reports, S&P Industry Surveys, Snapshots, TableBase, AB/INFORM, Business&Industry, Factiva, NewsEdge ... vilka jag räknar upp här bara för att visa spännvidden.

Alacra berättade bl a om att dess "flaggskepp" Alacra Book nu levererar den sammanställda informationen (om ett företag) även som MS Word-dokument (tidigare enbart i pdf). En helt ny tjänst är Alacra Corporate Connections, som kartlägger och redovisar personers befattningar och uppdrag i olika företag mm - och därmed klargör samband, "connections", mellan företag och "länkar" mellan deras befattningshavare.

Alacra startade 1996 (som Data Downlink). Antalet anställda är ca 30 i UK (plus ett 20-tal som anlitas externt) och ca 60 i USA (plus ca 50). Antalet kunder är ca 700. En antydning om vilken typ av kunder det rör sig om ges av att den person jag talade med i montern om ett par dagar skulle till Hong Kong - för att träffa en enda kund! Men så inbringar ju varje kund minst ett par tiotusen kr per månad.

Alacra har för övrigt under 2004 startat den nya databasen WorldData tillsammans med EcoWin och **EIU, Economist Intelligence Unit**, där den sistnämnda också ställde ut på Online 2004 (men som jag inte hann besöka). WorldData är en tjänst som i nästan realtid tillhandahåller ekonomisk information om regioner och branscher. Som bas nyttjas en databas med ekonomiska data (inkl förutsägelser/prognoser) om 150 länder mm, vilken kombineras med färskast möjliga uppgifter.

Apropå "primärproducenter" fanns flera producenter av marknadsrapporter närvarande på Online 2004, bl a **Snapshots** och **Euromonitor** (som också tillhandahåller GMID, Global Market Information Database, avsedd för bibliotek) samt förstås **MarketResearch.com**, som på ett bräde tillhandahåller marknadsrapporter från ca 350 sådana producenter. Ingen av dessa hann jag dock besöka, men att döma av pressmaterialet hade de ändå inga speciella nyheter att visa upp.

**Hemscott** tillhandahåller nu också information om företag och befattningshavare i USA och Kanada. Bakom ligger det förvärv man nyligen gjorde av amerikanska CoreData och dess databaser med denna information. Jag frågade när det var dags att utvidga till övriga Europa (Hemscott täcker ju f n enbart UK). Svaret var att man inte har några näraliggande planer på detta, "dock kanske med tiden". Svenska kunder? Inte som den jag talade med kände till i alla fall.

**Hoover's** var förstås med på Online, inte minst eftersom man nu åter satsar på Europa och har etablerat ett eget kontor i UK (vilket jag berättat om i nyhetsbrev 2004:6). Jag fick nu bekräftat det jag antagit, dvs att PowerInfo lagts ner och att dess personal (ca 10 personer) numera tillhör Hoover's. Trots detta fortsätter man (i den nya skepnaden alltså) att fungera som representant för såväl Skyminder som MarketResearch. Att Hoover's numera ägs av D&B syntes inte i montern. Men det märks på marknaden. Att D&Bs säljare numera marknadsför även Hoover's är en av orsakerna till Hoover's framgångar på senare tid.

Svenska **EuroInfoPool**, en mycken liten (men kanske prisbilligare?) aktör jämfört med konkurrenter som t ex Hoover's och Skyminder, ställde ut på Online för andra gången. Första gången var 2002, strax efter att tjänsten lanserats. EuroInfoPool är svenskt (ägs av InfoData) men man marknadsför sig huvudsakligen i UK. För att appellera till den publiken kallar man sig "The European House of Companies". Orsaken är förstås att UKs motsvarighet till vårt Bolagsverket heter Companies House. EuroInfoPool tillhandahåller ju en portal för samlad åtkomst till offentliga bolagsregister i 15 av Europas länder, kompletterad med kreditinformation om 25 miljoner företag (via Eurogate). En nyhet inför Online var en ny version av tjänsten.

**Moreover** har jag inte sett på Online tidigare. Företaget är känt som en av de främsta aktörerna vad gäller att monitorera webben för nyheter, artiklar etc från tusentals källor (i stil med svenska Agent25, Retriever m fl - fast större. I likhet med dessa har man inte heller avtal med källornas ägare). Moreover's specialitet har tidigare (i alla fall som jag känt dem) varit att betjäna sina kunder med nyheter inom lämpliga delområden för publicering på deras egna webbplatser. Nu har man flera tjänster, från vanlig bevakning/urval av nyheter (med leverans i nästan realtid) åt företag till en söktjänst, baserad på det insamlade "meta-materialet". Och antalet källor har vuxit snabbt. För ett år sedan samlade Moreover om jag minns rätt in material från sådär 3.000 källor på webben. Idag är antalet sådana källor uppe i ca 9.000 - och nu bevakar man också webbplatser, f n 750.000 st, och ett 40-tal nyhetskanaler. Materialet kan sorteras på 380 kategorier och levereras i XML. Moreover har ca 40 anställda. Totalt samlas 4 miljoner artiklar in per månad - och 3 miljarder (!) metadataposter levereras till kunderna. Bland partners finns Automomy och Verity - som nyttjar tillgång till information från Moreover som säljargument för sina söksystem. Bland svenska kunder/partners finns Akzo Nobel, Saab och Nobicon.

En de utställare jag hamnade hos utan att det var meningen var **DeHavilland**, UK. Dess tjänst DirectNews tillhandahåller nyheter anpassade till kundernas önskemål, avsedda att publiceras på deras webbplatser. Så långt inget märkvärdigt - men det lite ovanliga var att nyheterna inte tas från webben utan produceras/skrivs av journalister, anställda av DeHavilland (eller anlitate på timbasis). Det låter

dyrt - men uppenbarligen finns det en efterfrågan eftersom man har så många som 150 journalister anställda (Men man driver förstås också ett par andra tjänster, som PressBrief, PublicAffairsBriefing och AirWatch (nyheter för flygtransportsektorn i UK)).

## ***Övriga onlinetjänster, tidskriftsförmedlare, förlag m fl***

Några av de utställare utöver ovan nämnda jag kanske också borde besökt på Online var **EBSCO**, **Swets** och **ProQuest**, vilka jag dock prioriterade ner eftersom de ställt ut på vår egen (dvs SFIS/Informationsspecialisterna) höstkonferens i Stockholm veckan innan. Betr EBSCO och Swets tror jag det intressanta nya nämndes i nyhetsbrev 2004:6.

Men betr ProQuest fann jag i pressmaterialet en nyhet som jag noterat redan på SFIS-konferensen och som kan vara intressant att relatera här. Den är att onlinetjänsterna KnowEurope och KnowUK (vilka jag inte tror sålt särskilt bra) ersätts av den nya tjänsten European Sources Online, ESO. ESO baseras på referensskällor från Chadwyck-Healey. Därtill annonserade ProQuest att MLAs International Bibliography ska bli tillgänglig via Literature Online.

**Ingenta** presenterade inget nytt - utom att man nu ökat antalet publikationer man tillhandahåller till nära 29.000, varav 6.500 är e-tidskrifter. Så berättade man om det avtal man sedan i augusti har med Swets. (EBSCO har man avtal med sedan tidigare). Som ny vd efter grundaren Mark Rowse har nu utsetts Simon Dessain, tidigare driftchef. Antalet anställda tycks öka igen efter en period av nedskärningar och är nu ca 140. Ingenta tycks satsa en hel del på USA, där man också har egna kontor. I övrigt har man agenter bl a i Italien, Mellanöstern och Indien.

**Taylor & Francis** monter hamnade jag i utan avsikt - jag ser ju dem främst som förlag. Och så är det ju: T&F publicerar år 2005 mer än 1.000 tidskrifter och 2.000 nya böcker. Men nu fick jag i alla fall en inblick i en del av de referensverk som finns online och som marknadsförs under varumärket Routledge. Helt nya online bland dessa var Europa World Plus och Armed Conflict Database; sedan tidigare finns The International Who's Who, The World of Learning och Routledge Encyclopedia of Philosophy online. Men mest intressant var helt nya tjänsten Routledge Reference Resources. Den utgör ännu bara ett embryo, med bara två områdena tillgängliga: Religion och Politics & International Relations. Men tjänsten ska fyllas på med område efter område för att så småningom utgöra ett alltmer fullödigt "online library".

På utställningen fanns också förlag m fl, t ex Springer, Blackwells och John Wiley & Sons, tillhandahållare av (bl a) e-tidskrifter, t ex SAGE, HighWire Press och Infotrieve, och e-referensaktörer, t ex xrefer plus. Alla är de intressanta aktörer med "information", vilka jag dock nämner här (av sådär 100 ytterligare!) bara pga att de tillhandahåller sk presspacks - i vilka jag dock inte fann några större, helt nya, nyheter.

## ***Söksystem (enterprise-sökmotorer)***

Även om jag inriktade mig på information så blev det också ett par besök bland utställare med söksystem. Att undvika **FAST (FAST Search & Transfer)** var för övrigt i princip omöjligt för mig - delvis pga att "mitt" företag ju i höstas utsett FAST som ny leverantör av söksystem för bl a tjänsterna inom InfoTorg (vilket jag berättade i nyhetsbrev 2004:5) men framför allt pga FASTs köp av konkurrenten Nextpage i september (i samma nyhetsbrev). Det senare har ju också relevans för utvecklingen på marknaden för rättsinformation pga att Nextpage har en rad kunder just här, många av dem ägda av Wolters Kluwer (t ex Norstedts Juridik) (vilket jag berört ovan under Wolters Kluwer).

Alltså: vad händer nu med Nextpages installationer av NXT hos ca 800 kunder? Svaret var strax efter uppköpet (som vanligt vid sådana här uppköp) att den uppköpta partens produkter (dvs NXT) ska

fortsätta att existera. Men så brukar ju sällan bli fallet. Betr FAST finns ju fö exemplet AltaVista, vars söksystem köptes 2003 och det då sades ungefär samma sak. Nu tror jag att det är på väg att ersättas av FAST-installationer. Och när jag nu på Online frågade om NXTs framtid så var svaret (i min tolkning!) att FAST framdeles kommer erbjuda tre alternativ för befintliga kunder: 1) fortsätta med NXT, 2) migrera till FAST och 3) att nyttja FAST utan att lämna det befintliga gränssnittet (FAST bygger alltså ett "NXT-skal" ovanpå FAST, dit kundernas databaser migreras). Och alternativ 3) var ju precis det personen i Wolter Kluwers monter talade om. Framtiden får utvisa vad som händer, men det skulle förvåna mig mycket om FAST håller NXT under armarna även på sikt.

Företaget FAST tycks gå fortsatt bra, och nämns numera allmänt som ett av de tre stora inom området enterprise söksystem. De andra är Verity och Autonomy. Starkt marscherat på bara några år! Samtidigt utmanas ju alla tre av Googles nya Google Search Appliance, som ligger i samma prisläge men som har namnet och det goda ryktet att rida på - därtill något de andra inte har: närvaro på webben med egen sökmotor! Men lite finns FAST med på webbsökmotorområdet i alla fall: Scirus och Scopus bestyrkas ju av FAST.

Jag besökte också **Mondosofts** monter. Mondosoft är ett danskt företag med 20 anställda som utvecklar och marknadsför ett söksystem med samma namn. Starten skedde 1998. Huvudkontoret ligger i Köpenhamn och man har kontor även i Kalifornien, Norge och Frankrike samt ca 100 "strategiska partners". Även om Mondosoft inte tillhör dem som ligger i den absoluta toppen på söksystemsmarknaden hör man allt oftare dess namn; men att det har så många som 2.000 kunder hade jag inte anat. Mondosoft nyttjas för allt från webbplatser och portaler till intranät. Även om en fullfjädrat IR-sökmotor ingår så kan dess system för "live analytics" samt dess "BehavioralTracking"-system (som är ett verktyg för systemadministratörer att kolla användarnas beteende i syfte att förbättra sina tjänster) även nyttjas tillsammans med andra leverantörers sökmotorer.

Betr. leverantörer av söksystem/enterprise-sökmotorer fann jag i övrigt faktiskt bara Endeca. Autonomy, Verity och ZyLab, vilka "alltid" var med på Online tidigare har varit borta några år nu och tycks inte komma tillbaka. Franska nykomlingen Albert tycks ha lagts ner. OpenText minns jag inte har ställt ut någonsin (även om deras numera i LiveLink-serien integrerade BRS/Search och BASIS var med för länge sedan). Av "visualiseringssystemen" är bara anacubis kvar medan andra som ställde ut för några år sedan, som Wordmap, Lexiquest och OmniViz är uppköpta och/eller integrerade hos andra - och borta från Online.

## 5. Google Search Appliance

Google satsar nu på att bli en kraft att räkna med också på marknaden för enterprise-söksystem. Det gör man i och med en hårdansering av den nya, kraftigt förbättrade, versionen av Google Search Appliance, som kom i somras. Den första versionen kom 2002 och såldes främst i USA. Med den nya versionen satsar man lika mycket på Europa.



Som ett led i marknadsföringen bevistade Googles chef för enterprise-sökning, Dave Girouard, Online 2004. Där höll han en produktpresentation på utställningen samt ställde upp i (förbokade) intervjuer. Jag tog tillfället i akt och lyssnade dels på presentationen, dels intervjuade Dave. Referat resp intervju följer nedan. Men först vill jag förklara varför jag tar med så mycket om Google här i nyhetsbrevet, i synnerhet som en del av detta i princip kan betraktas som "reklam". Orsaken är att Googles förehavanden idag är av så stor betydelse för hela onlinebranschen att jag anser att det är av "allmänt intresse" att hålla sig ajour med dem.

## ***Produktpresentation: The business case for enterprise search***

Dave inledde produktpresentationen med att argumentera för vikten av att göra sökning såväl inom företag som sökning på deras webbplatser så enkel som möjligt. Han hävdade att "en så stor andel som en fjärdel av arbetstiden ägnas åt KM. En stor del av detta beror på sökverktyg som inte är användarvänliga". Vad gäller webbplatser så ger, enligt Jupiter, "85 % av dem inte svar på det som användarna söker efter".

Så avlivade Dave ett antal, som han sade, "myter". En sådan är att "författarna måste förse dokumenten med metadata för att möjliggöra effektiv återvinning". Faktum är, menade Dave, att ett sådant förfarande istället hämmar återvinning pga att många dokument därmed aldrig görs sökbara över huvudtaget. En annan myt är att "det är omöjligt att göra allt som finns tillgängligt i företagets nätverk enkelt sökbart. Faktum är att det är fullt möjligt att såväl fånga upp som att få det enhetligt och enkelt tillgängligt och sökbart via ett enda gränssnitt".

Lösningen heter "förstås" Google Search Appliance. Nyttja dokumenten på företagets nätverk som de är: metadata och annan förberedande behandling behövs inte! Allt som kan nås med http eller https kan hämtas och indexeras av Google, som gör dem enkelt sökbara på samma sätt som Google på webben. Och i och med att användarna redan är bekanta med denna så finns inte någon barriär mot utnyttjande. Det hela möjliggörs av Googles sökalgoritm, som dock här är något modifierad jämfört med den som nyttjas för webben: en del av dess över 100 faktorer är modifierade så att den bättre passar företag. Text är faktorn "antal till-länkar" nedtonad. Dave nämnde i sammanhanget att Google har över 1.000 personer sysselsatta med söksystemsutveckling, inkl arbete med sökalgoritmen.

Vad gäller kostnaden så är det visst så att Google Search Appliance kostar - men eftersom den innebär så lite arbete med installation och underhåll jämfört med andra söksystem så blir den mycket billig. Dave nämnde här att Världsbanken, som just köpt Google Search Appliance, bara kommer att behöva ägna en femtedel så lång tid åt underhåll jämfört med hittillsvarande söksystem. Ett annat företag behövde 9 personer för att sköta dess söksystem. Med Google klarar sig företaget nu med en person på halvtid.

Och vad gäller installation så kräver Google Search Appliance inte alls lika mycket insatser från IT-personal som andra produkter; samma sak vad gäller administration. En orsak till detta är att produkten säljs allt-i-ett - inklusive server och all programvara för drift och lagring. En enda box (eller multipla, beroende på företagets behov). Inget trassel med att fundera ut hårdvara och IT-miljö! Bara att plugga in på företagets nätverk och börja indexera! Såväl olika teckuppsättningar (stöd finns för 28 språk) som format (250 olika stöds) detekteras automatiskt.

För att inte tala om vilka resurser som spars genom bättre informationsåtervinning. Enligt Morgan Stanley har sökanvändningen ökat 11 gånger i företag som börjat nyttja Google Search Appliance - på grund av att användarna känner sig så hemma med sökgränssnittet. Dave påpekade att "Search should be a joy to use for employees and customers".

Men passar verkligen enterprise-sökning in i Googles "mission"? Jodå, sade Dave, det är precis i överensstämmelse med Googles inriktning på att lösa problem, inte på att tjäna pengar. Googles grundfilosofi kvarstår!

### *Åhörarnas frågor*

Vid den efterföljande frågestunden berättade Dave att en ny release av Google Search Appliance kommer 2005 men att den inte kommer att innebära någon extra kostnad för kunderna. Finns folk som kan bistå vid installation och informatik (typ förbereda information för sökning)? Nej, var det korta svaret. Vad kostar då kalaset? Ja, även där har man förenklat tillvaron för kunderna jämfört med konkurrenternas produkter. Köp en (eller flera) boxar, betala en klumpsumma - sedan får du använda



den bäst du vill, hur mycket du vill och till vad du vill, under två år. Uppgraderingar till ev. nya versioner ingår.

Någon frågade varför produkten inte levereras på CD istället för färdiginstallerad på en server. Svaret var att egen installation från en CD bara riskerar att leda till massa krångel. Nu har vi gjort allt klart "på fabriken" - bara att plugga in, indexera och köra. "Allt kan vara uppe och all information sökbar samma dag som boxen levereras". Hur är det med säkerhet, behörighetshantering, etc då?, undrade någon. Dave svarade att produkten stöder olika autenticeringsmetoder. Vad gäller behörighet till olika typer av material så kontrolleras sökresultat mot behörigheter på dokumentnivå. Kan sådan information som finns i företagets databaser (t ex DB2, Oracle etc) nyttjas? Dave svarade att en förutsättning är åtkomst via http eller https. Men arbete pågår för åtkomst även via andra åtkomstmetoder.

(Anm: I en artikel i The Register nämns en kanske oväntad "risk" med att indexera alla dokument i företaget med Google: den gör plötsligt alla dokument synliga, inkl sådana som ligger undangömda i nån mapp nånstans. Inte så kul om det t ex rör sig om företagets lönelistor!)

Mer info om Google Search Appliance på [www.google.com/enterprise](http://www.google.com/enterprise)

## ***Intervju med Dave Girouard, General Manager, Enterprise, Google***

### ***Funktionellt***

Jag inledde intervjun med funktionella frågor (som det anstår en informationsspecialist ...). Det är ju bra att det inte ställs krav på strukturerade dokument och metadata - men hur är det med funktioner för att tillvarata struktur och metadata *om* man vill? Finns fältsökning? Dave svarade att det ju inte är så vanligt att kunderna frågar efter det, men det finns i begränsad utsträckning. Stora dokument - är det OK (frågan ställd med tanke på t ex stora rättsfallsdokument)? Jadå, men det är klart att det finns *någon* gräns.

Google Search Appliance har stöd för booleska sökoperatörer som AND, OR och NOT samt användning av parenteser. Däremot finns inte proximity-(närhetssöknings)funktioner. Denna typ av funktioner finns dock implicit i sökalgoritmen/relevansfunktionen. I denna ingår även i övrigt samma funktioner som vid Google på webben (synonymer, stavningskontroll etc etc). (Anm: I vilken mån maskningsfunktioner/användning av wildcards finns missade jag att fråga. Jag förmodar att det inte finns explicit - men att även detta säkert nyttjas i sökalgoritmen.)

Allt material som går att "crawla" går att indexera. Men http och https i all ära: när kommer övriga möjligheter, typ att nyttja företagets databaser? Dave kunde inte ange annat än *att* det kommer.

### ***Priser/användning***

Googles prissättning är extremt enkel jämfört med konkurrenternas (vilka tillämpar en vild flora av varianter, allt från sökfrågor per sekund, antal indexerade dokument, antal servrar, antal processorer, antal användare, företagets storlek, antal applikationer, syfte (intern eller extern användning), mm!).

Så här ser prislistan ut:

Modell	Antal dokument	QPM (sökfrågor per minut)	Pris per två år
GB-1001	150.000	300	~ 250.000 kr
GB-1001 (alt)	1.5 miljoner	300	~ 1.4 miljoner kr
GB-5005	3 miljoner	300	~ 2.4 miljoner kr
GB-8008	15 miljoner	1.000	~ 4.7 miljoner kr



Detta gäller alltså två år. Inga avgifter för underhåll tillkommer under denna tid. Alla uppgraderingar ingår. Å andra sidan måste kunden köpa nytt/betala på nytt efter denna tid.

För den som kanske inte är så hemma på priser för söksystem kan sägas att de ligger i klass med vad de stora konkurrenterna tillämpar. (Se också fotot i anslutning till Steves Arnolds föredrag nedan.) Observera dock de stora skillnader som kan uppstå i och med de olika prissättningsprinciperna. För förhandlas nästan alltid priset fram. Nästan alltid är det också frågan om en initial licens som följs av en årlig underhållskostnad (vanligen kring 20 %). Dave hävdar att man aldrig ger rabatter eller förhandlar om priset. Prislistan gäller.

Ett sätt att delfinansiera ett köp kan vara att nyttja annonser på samma sätt, typ keyword match, (och i samarbete med, om jag förstod det rätt) som Google på webben.

Får Google Search Appliance nyttjas helt fritt/för att indexera vad som helst, dvs inte bara internt material utan även sådant som finns på nätet? Och får den nyttjas för den egna webbplatsen? Svaret är att man får indexera eget och licensierat material och förstås annat man har rätt att nyttja. Om det sedan bara nyttjas internt eller görs sökbart på webben är kundens ensak. Kostnaden för produkten är helt oberoende av hur den används - externt eller internt. Köp och kör!

### *Marknaden för enterprise-söksystem*

Vad säger då Dave om marknaden för enterprise-sökmotorer? För det första angav han att han är förvånad över att den är så liten, enligt hans egen uppskattning kanske 4 miljarder kr, i alla fall inte över 7 miljarder kr. "Den borde vara mycket större". Orsaken är, enligt Dave, att befintliga sökprodukter är alltför komplexa - i alla avseenden (installation, indexering, sökning, underhåll ...). Därför kommer Google att öka marknaden rejält!

Inte nog med att Googles produkt *är* enkel - varumärket är så bekant att den inte alls kommer att stöta på samma motstånd inom företag som hittillsvarande produkter (Anm: Ja, det är nog så sant som det är sagt. Om du säger Convera, Endeca eller Verity till en IT-människa så undrar de vad du talar om. Säg istället Google så vet de direkt!) I sammanhanget kan nämnas att "gamla" Google Search Appliance under 2003 omsatte drygt 400 miljoner kr - vilket inte är dåligt med tanke på att det var en så pass ny produkt. Orsaken står säker till stor del att finna i varumärket.

Men kommer då inte konkurrenterna på söksystemssidan att följa efter? "Jo, säkert - men det är bara bra!". (Anm: Jag noterade inte varför, men själv tror jag inte att det "drabbar" Google, som kommer att växa ändå. Därtill kan säkert spridningen av Google öka acceptansen för denna typ av söksystem - och alltså även hjälpa konkurrenter på traven - särskild sådana som tillhandahåller faciliteter som Google (ännu så länge ...) saknar och många företag faktiskt måste ha.

Hittills är Google det enda företag som har såväl en känd webbsökmotor som en enterprise-sökprodukt. (Anm: FAST lämnade ju webben 2003 i och att man då sålde AllTheWeb till Yahoo! - vilket man nog ågrar lite nu ...).

Här kom förstås osökt frågan upp om vad Yahoo! kommer att göra. Kommer de att följa efter? Nej, det trodde inte Dave. (Anm: Men nu i slutet av december har ju Yahoo! trots allt gett sig in på enterprise-sökmotorområdet på en lustig väg, nämligen genom en allians med Verity. Jag vet inte riktigt hur de tänkt sig det hela, men det är en annan historia som jag får återkomma till).

Hur ska Google Search Appliance säljas då? Jo, huvudsakligen av Google självt. Man etablerar nu säljteam i Europa, med början i Tyskland och UK. Dave upplyste om att produkten på bara en månad sålts i "ett dussin" exemplar - bara i Europa! (Anm: Det bör innebära kanske 30 miljoner kr - vilket är extremt mycket på så kort tid!). Sverige då? Svaret var här lite mer luddigt, men jag fick i alla fall

intryck att Googles nya säljkontor i Stockholm kan komma att nyttjas så småningom. Säljkontor - hur är det med teknikutstöd då? Nej, det ska inte behövas - i och med allt-i-ett-utförandet!

## 6. Föredrag

När Martin White inför konferensens inledande "keynote" presenterade Jakob Nielsen, webbens "king of usability", sade han att "nu kommer konferensens höjdpunkt"! Bara en artig fras, tänkte jag då. Men när konferensen var slut visade det sig att han hade rätt. Inget av de övriga föredrag jag åhörde var i alla fall såväl lika intressant som väl framfört.

### *Web usability*

**Jakob Nielsen**, Nielsen Norman Group, USA, har i över tio år arbetat med att göra webbplatser och intranät mer användarvänliga. Han har också författat flera böcker om webbdesign, bland dem "Designing Web Usability: the Practice of Simplicity", något av en "bibel" för webbdesigners och översatt till 22 språk, bland dem svenska ("Användbar webbdesign"). Den baseras på ett omfattande material från användarundersökningar.

Jakob talade nu över ämnet "*What's new, what's old in online usability*" - som också illustrerades just med några filmer från en användarundersökning. Den bestod i att ett antal försökspersoner fick olika uppgifter att lösa via 25 utvalda webbplatser. Uppgifterna kunde vara av typen "Du och din familj ska resa till Mexiko på semester. Sök och boka en lämplig resa", "Du vill investera 10.000 dollar i en pensionsfond. Leta upp det bästa sättet att investera dem." och "Du ska skriva en rapport om Marie Curie. Ta reda på vem hon är och vad hon är mest känd för". Filmerna visade försökspersoner i aktion, mot bakgrund av den skärmbild de arbetade på. Man kunde alltså följa hur de arbetade, t ex med musen.



Försökspersonerna berättade hela tiden vad de gjorde/försökte att göra. Det var oerhört illustrativt och man kände verkligen igen de "tabbar" som gjordes - beroende på dålig design. Mycket talande var t ex den situation där en försöksperson klickade på en länk för att komma till en bokningssida - men inget hände. Personen undrade varför och fortsatte att leta, kom till samma länk igen, klickade - men inget hände nu heller. Och så upprepades det hela tills personen gav upp. I själva verket hade hennes klick öppnat en hel rad nya fönster, vilka nu låg som knappar i aktivitetsfönstret.

Men denna demonstration tjänade bara som en introduktion till Jakobs val av tema, nämligen hur webbplatser och användare utvecklats med tiden. Tack vare sin långa verksamhet inom området kunde han relatera sina aktuella undersökningar till motsvarande äldre. I den ovan nämnda undersökningen lyckades försökspersonerna utföra uppgifterna i 65 % av fallen. Motsvarande frekvens år 1997 var bara 40 %. Slutsats: Det förväntade utfallet idag är att användarna lyckas med det de föresatt sig - när de väl har hittat fram till webbplatsen. Alltså är det detta senare som idag är det viktigaste vid webbplatsutveckling.

Jakob gick nu vidare med att redovisa vilka områden som är mest problematiska på dagens webbplatser. Så här stor andel av användarna upplever problem på resp område.

<i>Andel användare som upplever problem</i>	
<b>Finding (Info architecture, Category namnes, Navigation, Links)</b>	<b>35 %</b>
<b>Page design (Graphics etc)</b>	<b>25 %</b>
<b>Information (Content, Product info, Corporate info, Prices)</b>	<b>20 %</b>
<b>Search</b>	<b>15 %</b>
<b>Task support (Workflow, Privacy, Forms, Comparison, Inflexible)</b>	<b>15 %</b>
<b>Fancy design (Multimedia, Back button, PDF/Printing, New window, Sound)</b>	<b>10 %</b>
<b>Other (Bugs, Presence on Web, Ads, New site, Metaphors)</b>	<b>5 %</b>

Det är alltså sådant som har att göra med webbplatsens struktur som oftast upplevs som problematiskt medan sökning är enklare. "Fancy design", ett av de största problemområdena på 1990-talet (bl a pga att många webbdesigner experimenterade med massa finesser) möter numera inga större problem.

Vad gäller webbanvändning generellt så är det första användarna företar sig i 88 % av fallen att nyttja en webbsökmotor (typ Google). För varje "uppgift" som användarna företar sig på webben besöks i genomsnitt 3.2 webbplatser (exklusive webbsökmotor) - och i snitt 0.4 av dessa besöks ytterligare en gång (eller mer). I 40 % av fallen är det webbplatsens förstasida som besöks först - och då i 35 sekunder (ovana användare) eller 25 sekunder (erfarna användare). I de 60 % av fallen som användaren först kommer in på en "intern" sida så spenderas 60 sekunder (ovana) eller 45 sekunder (vana) på denna. Slutsatsen är att man har väldigt kort tid på sig att fånga användarens uppmärksamhet. Alltså är det extremt noga med en bra utformning på de sidor som man vet att användarna först kommer till.

I snitt besöker användarna 3.3 sidor på en webbplats i de fall de inte finner vad de söker och går vidare till annan webbplats - och tiden som spenderas är 1 minut och 49 sekunder. I de fall det rör sig om den sista webbplatsen för den aktuella uppgiften så besöks 5.5 sidor - och spenderad tid är 3 minuter och 49 sekunder.

#### *Användarna blir allt duktigare - använder allt fler sökord*

Väldigt intressant tyckte jag det var att ta del av Jakobs redovisning av vilken förändring som skett med tiden vad gäller det antal sökord som användarna nyttjar vid sökning i sökmotorer:

	<b>1994</b>	<b>1997</b>	<b>2004</b>
<b>1 sökord</b>	<b>81 %</b>	<b>44 %</b>	<b>36 %</b>
<b>2 sökord</b>	<b>14 %</b>	<b>33 %</b>	<b>36 %</b>
<b>3 sökord</b>	<b>4 %</b>	<b>14 %</b>	<b>16 %</b>
<b>4 sökord</b>	<b>1 %</b>	<b>5 %</b>	<b>6 %</b>
<b>&gt; 5 sökord</b>	<b>0 %</b>	<b>2 %</b>	<b>7 %</b>
<b>Antal i snitt</b>	<b>1.3</b>	<b>1.9</b>	<b>2.2</b>

En stor förändring har alltså skett på dessa tio år. Användarna har helt enkelt blivit mycket duktigare, sökmotorerna blivit bättre och webben blivit allt större. Frågan är nu om trenden kommer att bestå? Ja, det trodde Jakob, som förutspådde en utveckling mot närmare 3 sökord i snitt per sökning. Egentligen är det ju inte alls mycket mot bakgrund av de många miljarder sidor som finns sökbara i de stora sökmotorerna!

Finner då användarna vad de söker? Ja, i 42 % av fallen gör de det (men det skiljer mycket mellan ovana användare, som lyckas i 32 % av fallen, och de mest erfarna, 50 %). En intressant uppgift är att 48 % av användarna bara tittar på sökresultatets första svarsida (dvs de bläddrar inte vidare). Och sökningar på webbens sökmotorer lyckas i 56 % av fallen - men bara i 36 % av fallen finner användarna vad de söker när de anlitar webbplatser.

Söksidans utformning är mycket viktig vad gäller att se till att användarna får bra sökresultat. Jakob rekommenderade pedagogiska sökexempel som ett sätt att bra sätt att nå dithän.

Jakob har funnit att i följande avseenden upplevs samma problem som på 1990-talet:

- **Besökta länkar ändrar inte färg**
- **Bakåtknappen fungerar inte i alla lägen**
- **Nya webbfönster ("hidden windows")**
- **Pop-up-fönster**
- **Ser ut som reklam**
- **Brott mot vedertagna webbregrer/webbstandarder**
- **Innehåll som inte säger något**
- **Alltför kompakt text; text som inte går att snabbläsa**

Viss bättring har skett vad gäller bland annat sidor där utskriven text inte ryms på sidan, sidor där man inte ser vad som är klickbart, fryst layout (t ex text) som borde vara "flytande" och krav på registrering. Avsevärda förbättringar har skett vad gäller tid för nedladdning, användning av frames, användning av symboler som man inte vet vad de betyder, webbdesigners som experimenterar för mycket med "det allra senaste", standardiserad design, reklam-logotyper som uppfattas som webbplatsägarens logotyp, upplysning om vem som står bakom webbplatsen, m fl. Av de totalt 34 kriterier/regler för design som Jakob nyttjat sedan 1994 gäller sammanfattningsvis idag 2004 följande:

- 8 kriterier (se ovan) Fortfarande problem. Väsentligt att åtgärda
- 10 kriterier Fortfarande problem. Av medelstor betydelse att åtgärda
- 14 kriterier Fortfarande problem. Av liten betydelse
- 2 kriterier Inte längre något problem

Jakob gav slutligen en liten ovetenskaplig prognos. Det är 34 procentenheter kvar till maximal användarvänlighet. Användarvänligheten på webben har hittills ökat med i snitt 4 procentenheter per år. Alltså tar det 8-9 år tills idealtillståndet uppnåtts! Men han avslutade med att säga att vad som är god design var känt redan när webben slog igenom. Studie efter studie, år efter år, visar bara det vi egentligen redan vet!

## ***Search: is there a life after Google?***

Moderator för denna session, med tre föredrag, var välkände Chris Sherman, SearchEngineWatch, en av författarna till boken The Invisible Web. Han tog sig friheten att omedelbart svara på frågan i sessionens rubrik: "Ja, det finns ett liv efter Google!" Så lämnade han över till Steve Arnold för sessionens "keynote".

## ***Search: Outlook 2005***

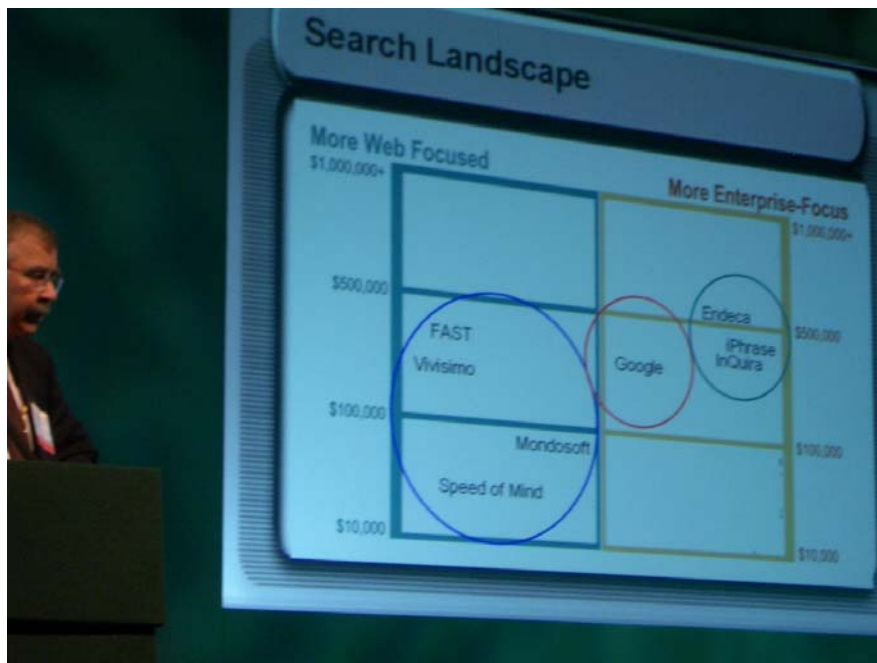
**Steve Arnold**, Arnold Information Technology, USA, som är en guru vad gäller sökmotorer/söksystem, inledde med att påpeka att "enterprise search" (ES) är något helt annat än "webbsökning". Den viktigaste skillnaden är de krav på säkerhet/behörigheter som finns vad gäller ES men inte för webbsökning. Han gjorde också "reklam" för sin nyligen publicerade bok "Enterprise Search Report", som bl a innehåller beskrivning och analys av 26 av marknadens ES-system, bl a Atomz, Automony, Convera, dtSearch, Copernic, Endeca, FAST, Google Search Appliance, iPhrase, Open Text/Livelink, Mondosearch, Verity och Vivisimo Clustering Engine samt ger tips för utvärdering och val av söksystem.



Hans föredrag fick en extra krydda av att den nya versionen av Google Search Appliance nyligen lanserats; ett produkt som han menade kommer att påverka ES-marknaden mer än något annat under 2005. Google på webben tjänar som ett "lockbete" för Google Search Appliance - vilket kommer att ge den en ordentlig skjuts. Så gav han lite fakta kring produkten - samtidigt som han förutsade att sådana avancerade funktioner som den inte har idag men som finns hos konkurrenterna helt klart kommer att utvecklas på sikt.

Steve åskådliggjorde ES-marknaden i ett par diagram, varav ett framgår av fotot. Han delade in

"before-Google" systemen i två kluster, "Web Centric Search Systems" respektive "Enterprise Search Systems", varpå han till den första gruppen (som framgår av fotot) hänföde FAST, Vivismo, Mondosoft och Speed-of-Mind och till den senare Endeca, iPhrase och InQuira. (Anm: Jag håller inte riktigt med om denna indelning, men det är ju Steves föredrag ...). Så menade han att Google nu tar det tomma utrymmet mittemellan: hyfsat prisläge - och passar hyfsat för både webb och enterprise.



Att Google kommer att få framgång med sitt ES-system beror också på att företagen ännu inte fått upp ögonen för att det är mycket mer än sina databaser som måste göras sökbara på sikt. Databaser utgör i snitt bara 23 % av all information. Det övriga utgörs av filer (21 %), bilder (16 %), pdf (12 %), övrigt elektroniskt (5 %) - och pappersbaserad information (23 %). Och många av marknadens ES-system som utger sig för att klara "allt" är när det kommer till kritan inte anpassade till verkligheten.

Steve menade att genombrottet för ES ännu inte kommit. Det pågår en rad projekt baserade på innovationer inom sökning, bl a Attensity, DocumentFinder (GetPDF.com), EOS Web Indexer, Find It Pro (Skylark), Masala (IBM), Microsoft Search, Petamem, PowerSearch (Hnisoft), Search 32 (Anet), SliDR (NiSUS), Smart Miner (ServiceWare), Sourceware Search (Market Central), Spy CD Search (PHDCC.com) och Subject Search Sleuth (Kryloff) - men av dessa finns det bara två som har en chans att hävda sig, medan de andra kommer att få det minst sagt tufft på marknaden för ES-system.

Den ena av de två med chans är Masala (IBM), som är byggd på WebFountain (Anm: Se också ovan under Factiva). Masala finns för ut i en första version nu, under namnet OmniFind.). Den andra är Microsoft Search, men den kommer att komma först 2007 (den är en del i Microsofts projekt Longhorn). Som kommande trender angav Steve dels visualisering (alla Inxight), dels sökning baserad på och automatiskt anpassad till användarens beteende.

Slutligen gav Steve rådet till dem som står i begrepp att utreda ES-system att inse att söksystem och indexering bara utgör en mindre del av totalkostnaden för anskaffningen. De största kostnadsposterna utgörs i själva verket av att förbereda data (nytt format, konverteringar mm), tillgodose önskad prestanda (servrar, kommunikation, etc) och "handpåläggning"/manuella åtgärder av olika slag (design, kategorisering, klassificering, mm).

### ***The new frontier of search: using advanced search techniques for improved decision making and global information perspective***

En lång titel på ett kort föredrag bestod oss **Bobbie Browning**, FAST (USA-kontoret) med. Nej, förresten, inte kort men däremot ganska innehållslöst, i alla fall att döma av mina anteckningar - och bara en halv sida i proceedings. Därtill bestod det i alltför hög grad av marknadsföring av FAST, enligt min mening.

Bobbie inledde med att konstatera att dagens "key driver" är sökning med utnyttjande av en enda sök-ruta (dvs det jag ovan kallat "Googlemania") och att 70-80 % av all information som behöver sökas är ostrukturerad. Vad gäller det senare är nästan all information vid myndigheter av den karaktären medan banker ligger i topp vad gäller hårt strukturerad information. Av dagens företag med ES-system är det få som bara har ett enda sådant: de flesta, 64 %, har 2-3 st, medan 21 % har 4-5 st. Många skulle kunna ersättas av ett som kan indexera allt. Så gav hon några exempel på sådana installationer (Anm: försök gissa vilket ES-system de nyttjar ..). Det första var FirstGov.gov, dvs USAs motsvarighet till sverige.se. Det andra var Law Library of Congress GLIN (Global Legal Information Network) vars tidigare använda system Inktomi inte kunde växa tillräckligt för att klara kraven. Det tredje var jordbruksmyndigheterna i USA, som monitorerar och indexerar ca 36.000 webbplatser.

### ***The Semantic Web and Knowledge Management***

Sessionens sista föredrag hölls av **John Davies**, BT Research & Venturing (UK). Han började med att "undervisa" oss i XML, ontologier och RDF. Så kom han in på sitt huvudämne: Projektet SEKT (Semantically-Enabled Knowledge Technologies), vilket adresserar de krav som ställs av de kommande NGKM-systemen (Next Generation Knowledge Management). SEKT är mycket stort: budgeten är på ca 12 miljoner euro. Det stöds av EU-kommissionen och drivs av 12 partners. Ambitionerna är att förbättra automatisk ontologi- och metadatagenerering, utveckla lösningar som är extremt skalbara, forska kring hur inkonsistent material kan nyttjas samordnat, utveckla verktyg för semantisk KM och utveckla metodologi för hur verktygen ska nyttjas.

Den inressantaste delen av föredraget redovisade exempel på analys av stora mängder dokument och texter med användning av semantiska och andra hjälpmedel, allt i syfte att automatiskt identifiera och märka upp olika entiteter, t ex personer och företag, i dessa.

John listade så det han ser som trender vad gäller sökmotorer. De fem viktigaste sådana rör "desktop search", kategorisering, integrerad sökning (samtidig sökning i olika typer av material), "seamless search" (sökssystem inbyggda i de ordinarie KM-systemen) och "personalised search".

Han avslutade med att konstatera att de förbättringar som kan uppnås vad gäller rankning av sökrultat är högst marginella. Därför måste nya metoder nyttjas. Två ansatser finns, dels "den semantiska webben" (SEKT är en representant för denna), dels "Beyond search" (representant bl a Corpora, som dock inte beskrevs). Till detta kommer nyttjandet av lingvistiska metoder och ontologier (för att "mäkla" mellan olika vokabulärer). Framtiden är "information-centric, not document-centric!"

### ***Can you trust what you find?***

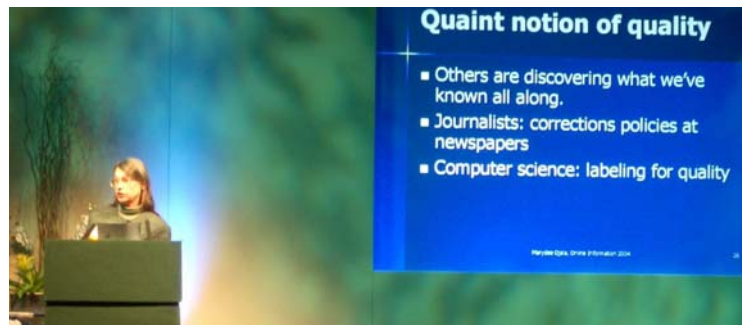
#### ***The quaint notion of quality: credibility meets credulity.***

I denna session lyssnade jag på två föredrag. Det första hölls av välkända **Marydee Ojala**, chefredaktör för Online Magazine (USA) och hade titeln "The quaint notion of quality: credibility meets credulity". Marydee är väldigt duktig kunskapsmässig och en bra föredragshållare. Jag blev ändå besviken.

Orsaken var att hon höll sig på en väldigt basal nivå. Ett exempel är att hon till och med beskrev vad "recall" och "precision" är. Och råden hon gav var så självklara.

Marydee gav exempel på vad man ska vara på sin vakt mot på webben: Falska webbplatser (dvs sådana som ser ut som riktiga, existerande, men i själva verket är nästan exakta kopior), konstiga sökmotorresultat, vilseledande uppgifter (som ser ut som äkta), källor som inte är kompletta, falska historier/myter, tekniska fel (som kan förvränga resultat). Varning också för att funnen information kanske inte alls är sann - det som ser pålitligt ut kan mycket väl vara en hobbysida. Därtill har felaktig information på webben en tendens att "fortplanta" sig. Det som står på en webbplats kanske nyttjas av andra - och så småningom blir det "sanning" bara för att det förekommer på så många ställen. Marydee pekade också på de spratt som felaktig stavning kan ge - vad gäller sökning såväl som i svaren (Thomson - Thompson).

Man ska också vara medveten om att mycket av den information som finns på webben finns där i syfte att marknadsföra ett företag, produkt, ideologi, etc etc. Det innebär givetvis stor risk för att sådan information kan vara "biased". Därtill finns mycket som Marydee kallade "junk science". Slutbilden hon visade framgår av fotot här intill.



### *Addressing the data quality challenge*

Nästa föredrag i sessionen hölls av **Beth Jarusco**, OneSource, USA. Det hade i konferensprogrammet titeln "Addressing the data quality challenge" och i proceedings "OneSource and content quality". Även detta var på en basal nivå. Beth beskrev hur viktigt det är att den information som används i ett företag har bra kvalitet och vilka förluster dålig information kan orsaka. Hon redogjorde också bl a för informationens "Accuracy", "Timeliness", "Completeness" och "Relevance". Inte särskilt givande.

### *Your searching questions answered - The Eureka! Forum*

Förra årets Eureka Forum var en av de mest givande sessionerna under konferensen. Därför var mina förväntningar nu höga. Tyvärr kom de en aning på skam. Det kändes som om panelen, bestående av "sökkändisarna" **Gary Price**, **Karen Blakeman**, **Mary Ellen Bates** och **Chris Sherman**, flankerad av moderatorn Phil Bradley, var lite trötta. Det blev liksom ingen fart, inga skiljaktliga åsikter. Samtliga har ju figurerat på konferens efter konferens i år: kanske hade de fått frågorna som nu ställdes flera gånger förut? Nåväl, här kommer i alla fall en del av mina noteringar. En del intressant blev det ju trots allt.



Inte oväntat vilade Googles ande över hela sessionen - men samtidigt gick mycket i samtalen ut på att komma bort från Google. Dock handlade nästan allt om webbsökning. Chris inledde med att uppmana oss att gå in för att pröva "serious" alternativ till Google. Varför inte Yahoo!, som ju försågs med en ny sökmotor i mars, eller helt nya MSN Search (som kommer att bli bättre och bättre)! Det finns många andra som är värda att provas, t ex Clusty.com. Mary Ellen

kompletterade med att påpeka att MSN Search tillhandahåller en rad trevligheter som hon rekommenderade oss att titta på. Vad sägs t ex om möjligheten att med hjälp av dragreglage själv ställa in vilken vikt olika variabler ska ha vid sökning och relevansbedömning. Så tillade hon: "Tänk om man kunde få denna möjlighet i onlinetjänsterna!"

Karen instämde betr Yahoo!, som hon menade är en "fantastic search engine". Hon tipsade också om Find.com (en sökmotor inriktad på affärsinformation). Gary avslutade denna första omgång med talarna genom att instämma såväl betr tipset om dragreglage som Yahoo! men också genom att tipsa om Exalead, en fransk sökmotor som han menade passar bra för generella sökningar.

#### *Vilken roll får Google Scholar?*

Någon i publiken undrade vad panelen ansåg om Google Scholar. Karen menade att den är intressant men att den har stora brister; t ex är den dålig betr fältsökning (t ex behandling av författarnamns olika stavning). Gary menade att Google ska ha all heder i världen för denna tjänst; en tjänst som han ändå inte trodde skulle hota de kommersiella onlinetjänsterna. Chris berättade att Google Scholar är en frukt av Googles princip att alla Google-anställda ska ägna 20 % av sin arbetstid åt egna projekt (vilka kan baseras på vilka idéer som helst). Mary Ellen sade att tjänsten verkligen är bra för biblioteken; det är ingen leksak utan en seriös tjänst: Google har licensavtal med en lång rad förlag, institutioner etc för tjänstens innehåll. Det innebär också att en del av den fulltext som det länkas till och som inte är gratis faktiskt kan återfinnas gratis på annat håll, t ex vid universitetens webbplatser. (I sammanhanget tipsade Gary om Smealsearch, en sökmotor för pdf-dokument inom affärsinformation mm).

Hur kan då onlinetjänster typ Dialog komma att påverkas av Google Scholar? Gary svarade att de måste visa upp sina produkter bättre än idag - de måste tala om hur mycket licensierat, på andra håll avgiftsbelagt, material som faktiskt ingår i abonnemangen på deras tjänster. Chris relaterade för sin del marknadsanalytiker som tror att Google Scholar och andra kommande tjänster hel kan förstöra onlineindustrin. Mary Ellen pekade dock på det positiva i att sådana som aldrig skulle komma på tanken att nyttja en kommersiell onlinetjänst nu upptäcker vad som finns. Diskussionen om Google Scholar avslutades med att Gary påpekade att tjänsten inte är fullt så seriös som det fina ordet "scholar" antyder. Faktiskt är innehållet en rätt blandad kompot.

#### *Klustring, visualisering, specialiserade resurser, desktop search*

En publikröst frågade nu vad panelen trodde om framtiden för klustring. Tja, se bara på Clusty, sade Gary. Den används redan idag för annat än webbplatser/webbsidor, t ex för att "klustra" produkter från olika e-handelswebbplatser. Mary Ellen sade att hon använder Clusty när hon inte hittar det hon söker på andra webbplatser. Eftersom det är Vivisimo ligger bakom Clusty så kom diskussionen in på visualiseringsverktyg, t ex Grokker. Chris, som verkligen "vill" gilla dem, ansåg dem dock ännu inte tillräckligt utvecklade för att vara användbara. Varpå Gary inflikade att det vore kul att visualisera resultat från Google Scholar.

Hur arbetar då paneldeltagarna? Mary Ellen sade att hon istället för att leta efter "svar" på frågor letar efter "dörrar att öppna". I synnerhet använder hon bloggar. Karen använder ofta "format" (t ex pdf eller spreadsheets) som nyckel för att sälla bland sökresultat. Gary använder ofta specialiserade resurser som RDN (Resource Discovery Network, [www.rdn.ac.uk](http://www.rdn.ac.uk)), Globalspec (The Engineering Search Engine) och Infomine (Scholarly Internet Resource Collections, [infomine.ucr.edu](http://infomine.ucr.edu)). Chris berättade att han istället för att spara länkar till funnet material ofta tar ner det till den egna hårddisken - för att senare söka upp det där. Orsaken är att man inte kan lita på att materialet finns kvar (eller kvar på samma plats) på webben, som ju rör sig hela tiden.

Med anledning av detta ombads publiken att ange i vilken mån de nyttjar "desktop search" (indexering och sökning på den egna hårddisken). Bara sådär 5 % räckte upp handen! Chris varnade i sammanhanget för att innehållet potentiellt exponeras för hela det nätverk som den egna arbetsstationen är ansluten till. Karen berättade att hon tyckte lokal sökning fungerade dåligt och helt enkelt inte värt



tiden det tar att jobba med indexering etc. Gary jämförde lokala sökverktyg från Google och Copernic: "Fördelen med Google Desktop är att den 'tar allt' ". Den enda fördelen med Copernic är att den är lättare att använda." Gary nämnde också X1, den produkt på vilken Yahoo! baserar sitt lokala sökverktyg. Mary Ellen uppgav att hon för sin del verkligen gillar Google Desktop.

Vilken var den "trevligaste" sökmotornyheten under det gångna året? undrade nu Phil. Gary svarade nya Yahoo! news search samt topix.net (en "wonderful" resurs som han använder väldigt mycket), Mary Ellen svarade A9.com och Chris svarade Yahoo! directory i ny tappning. AltaVista då, undrade Phil vidare. Nej, svarade Chris, den finns det absolut ingen anledning att nyttja längre.

### *Framtiden?*

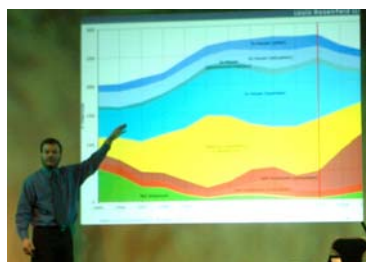
Hur är det med Google och framtiden? Chris upplever att det är ett skifte på gång. Allt fler "Google-killers" är på väg att utvecklas. Därtill finns det helt enkelt ett begynnande motstånd mot att nyttja Google. Mary Ellen hade under en kurs gjort en liten enkät bland folkbibliotekarier - och det visade sig att 25 % av dem undvek att nyttja Google. Karen uppmanar alltid sina kursdeltagare att inte saklöst använda Google: "det finns alternativ!".

Mary Ellen påpekade också det omöjliga i att lägga all information i en enda jättestor "låda", skaka den och tro att man kan göra justa sökningar i den. "Jämför matlagning! Man lägger inte allt i en kastrull och rör om!" "Och så undrar jag hur mycket tid vi förlorar på att göra på detta sätt istället för att leta på rätt ställen", tillade Gary. Därtill är det ju så väldigt mycket som inte indexeras. Google når absolut inte allt på webbplatserna. För att inte tala om "den osynliga webben" - som är minst 250 gånger så stor som den "synliga" (den som indexeras av sökmotorerna). För att inte tala om allt äldre material. Och även om allt detta skulle indexeras: är det verkligen bra? Tänk på den fara som ligger i att det bara är de 4-5 första träffarna på söklistorna som användarna "ser", sade Gary, som menade att vertikala (t ex ämnesspecifika) sökertyg kommer att bli allt viktigare. Frågan är bara hur användarna ska hitta dem!

## ***Trends in information architecture***

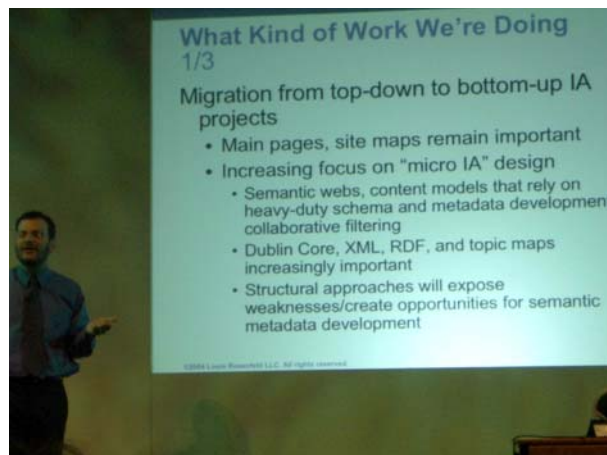
**Louis Rosenfeld**, oberoende IA-konsult (Louis Rosenfeld Association, USA), höll keynote-föredraget "Trends in information architecture". Han inledde med att göra reklam för den bok han skrivit tillsammans med Peter Morville, "Information Architecture for the World Wide Web", som kom 1998 men nu finns i en ordentligt utvidgad upplaga. Apropos 1998 så menade Louis att det inte alls är sådan "uppståndelse" kring IA nu som då - och IA-folket har blivit "civiliserat". Samtidigt kunde han konstatera att det egentligen inte hänt så mycket under denna tid; det råder en "lack of innovation". (Anm: Jämför med Jakob Nielsens liknande syn vad gäller utvecklingen inom området webbdesign!).

Louis ägnade största delen av sitt föredrag åt att redovisa en studie av trender inom IA under de senaste tio åren. Studien baserades på svar från 287 personer verksamma med IA-uppgifter. Frågorna avsåg IA-arbete under vart och ett av åren under perioden 1995-2004 samt förutsägelser för åren 2005, 2010 och 2015. Två huvudfrågor ställdes: 1) Vilken typ av arbetsplats arbetade du vid? och 2) Hur stor andel av din arbetstid ägnade du åt IA? Louis redovisade svaren i form av diagram - vilka jag är frestad att inkludera här men inte vågar, så det får räcka med ett par suddiga foton medan du själv måste klicka på [www.louisrosenfeld.com/home/bloug\\_archive/images/041017a.jpg](http://www.louisrosenfeld.com/home/bloug_archive/images/041017a.jpg) respektive [www.louisrosenfeld.com/home/bloug\\_archive/images/041017b.jpg](http://www.louisrosenfeld.com/home/bloug_archive/images/041017b.jpg) för att se detaljerna!





Diagrammen talar delvis för sig själva så jag relaterar inte Louis kommentarer kring dem (en del framgår f ö av hans egen beskrivning på [louisrosenfeld.com/home/bloug\\_archive/000307.html](http://louisrosenfeld.com/home/bloug_archive/000307.html)). Här



bredvid ses Louis i full aktion under sin redovisning (och där framgår ju också en del av detta ...).

"Vilka är vi då, vi IA-människor?" undrade Louis. "Handelsvara" eller en specialiserad resurs? Både/och - i symbios! Generalister eller specialister? Tja, båda områdena växer, men generalisterna utvecklas till strateger. Integratörer? Ja! (t ex avseende sådant som sökning, taxonomier, navigation). Informationsterapeuter? Ja, det kanske är det bästa sättet att beskriva vad vi kommer att arbeta med i framtiden!

Louis avslutade med att ange två avgörande skäl för att vara optimist vad gäller IA-arbete framgent: 1) Informationsexplosionen, 2) "information ROT (Redundant, Outdated, and Trivial content)".

## *De-mystifying enterprise search*

**John Lervik**, vd FAST (Fast Search & Transfer), Norge, inledde denna session med sitt föredrag benämnt *The new perspective on enterprise search*. FAST startade 1997 och är nu, enligt John, det näst största söksystems företaget med över 1.500 kunder. Han började med att konstatera att det nu var jämnt 5 år sedan han sist stod här i talarstolen.

Jodå, det minns jag mycket väl. Det var alltså på Online 1999, och så här skrev jag i min rapport, tillika det allra första numret av Online- och webbnyheter, 1999-12-28:

*"Hur skulle nu **John Lervik** från norska Alltheweb klara sig efter dessa herrar \*)? Inga problem! John berättade om sökmotorn FAST, som företaget utvecklat och som är en av de mest kända på webben. Arkitekturen ger obegränsad skalbarhet, åtminstone upp till en miljard dokument, och den kan växa allteftersom nya användare strömmar till. Den har "oöverträffad pris/prestanda" och medger nyttjande av lågpris-PCs. kan nyttjas. Med tanke på vad David \*) sagt sade John att de alltid uppger ärliga siffror om storlek etc. Idag hänvisar man till 210 miljoner URLs och man kan indexera 300 miljoner per dag. I vår ska man ha 300 miljoner dokument och i början av hösten 600 miljoner."*

\*) Här syftas på David Seuss från dåvarande Northern Light, som senare blev uppköpt av divine, snart därefter i konkurs, samt Mike Gallagher från Powerize.com, senare uppköpt av Hoover's. Så kan det gå!

Ja, det är mycket som hänt sedan 1999! T ex har FAST gått från rent webbsökmotorföretag till att bli ett rent söksystems företag. (Anm: Något som Steve Arnold, relaterad ovan, menade att FAST nu bör ångra lite i och med att man därvid avhände sig en extremt bra marknadsföringskanal).

I likhet med den tidigare föredragshållaren från FAST, Bobbie Browning, kunde inte heller John låta bli att göra lite väl mycket reklam för sin Enterprise Search Platform (ESP) - må så vara att den här fick tjäna som en exponent för det som John framställde som den ideala produkten; ett system som klarar alla typer av information.

Centrala egenskaper hos söksystem är flexibilitet, skalbarhet, tillgänglighet, säkerhet, "analysförmåga", pålitlighet/relevans. Här pekade John speciellt på analys, som han angav som en av de viktigaste

orsakerna bakom Googles framgångar. Hans lista över viktiga faktorer vid val av söksystem framgår av fotot nedan.



Därefter exemplifierades det hela med några exempel - såklart från installationer av FAST, bland dem Scirus och Scopus eftersom de tillämpar just analysmöjligheten (som i dessa fall delar upp sökresultatet i kategorier med möjlighet för användaren att med hjälp av dessa "förfinas" sin sökning).

John avslutade med att ange att en viktig trend den närmaste framtiden är söksystemens förmåga till "personalisering". Detta är speciellt viktigt för webbens söktjänster (t ex Google, Yahoo!, MSN Search). Orsaken är att man därmed knyter användarna/kunderna närmare till sig.

### *Choosing the perfect search engine for your organisation*

Så följde **Mikael Thorson**, InfoSphere AB (Sverige) med föredraget "Choosing the perfect search engine for your organisation". Det bestod till stor del av en redovisning av den metod som han och hans företag (som främst arbetar med konsulting) nyttjar när de hjälper sina uppdragsgivare att välja söksystem. Som vanligt framhävde Mikael vikten att se till att det system som väljs kan hantera ostrukturerad information, till skillnad då mot strukturerad information, som redan är väl täckt vad gäller system.

Mikael angav följande tre grundstenar (och åtgärder) att utgå från vid val av system:

- Användare: Intervjua representanter för avsedda användargrupper
- Sammanhang ("context"): Granska befintlig information
- Teknik: Studera organisationens tekniska miljö och dess förutsättningar att klara krav på prestanda

Men egenskaper i all ära - det kanske viktigaste vid val av söksystem är ändå den tänkta leverantörens stabilitet och överlevnadsförmåga, menade Mikael, som angav följande trender på marknaden för söksystem:

- Många små aktörer primärt fokuserade på USA och ganska ointresserade av Europa
- Viss konsolidering är på gång - större, etablerade aktörer förvärvar teknikföretag som kan förse deras produkter med analysverktyg
- Omvärldsbevakning, säkerhet och försvarsområdet är heta områden
- Många produkter har fortfarande dåligt språkstöd
- Många system som ser väldigt bra ut "på pappret" har svårt att nå upp till sina löften när de väl är på plats (Anm: jämför Steve Arnolds liknande varning)

### *Searching for the future: research trends in search*

**Steve Whittaker**, Information Studies Department, Univ. of Sheffield, UK, höll ett föredrag som i alla fall för mig var betydligt mer givande än de andra i denna session. Orsaken var att det gav konkreta beskrivningar, delvis i form av exempel, på vad framtida (men redan under utveckling) söksystem förmår. Steve angav fem viktiga områden: Gränssnitt, Multimedia, "Multilingual research", Geografiska system och Personalisering.

Steve ägnade merparten av sitt föredrag åt multimedia, där han bl a angav intressanta (men möjligen lite obehagliga ...) exempel på tillämpning av tal-till-textteknik. I och med framstegen inom detta område är det möjligt att med hyfsad kvalitet skapa text av tal (andelen fel är nere på 16 %, vilket är fullt tillräckligt för denna typ av tillämpningar), vilken sedan dels kan analyseras automatiskt för att finna samband, dels kan indexeras för att kunna söka fram intressanta delar av stora textmängder. (Kan också ske i realtid). Detta kan t ex nyttjas för att analysera de gigantiska mängder inspelade telefonsamtal som lagras av t ex terroristbekämpande myndigheter.

Ett annat exempel är att en kundtjänst med automatik kan göra sammanställningar av stora mängder telefonsamtal för att spåra områden med hög problemfrekvens utan att operatörerna ständigt behöver ägna tid åt att skriftligt dokumentera sina samtal. Det passar även bra för analys av samtal som kommer in på icke manuell telefontid och spelas in.

Språkforskningen är idag så långt framme att dess resultat kan nyttjas för t ex system där användaren utan vidare kan nyttja ett visst språk för att söka i material på andra språk - och också få svar på det egna språket (det senare har ju tillämpats i flera år nu, t ex i Google).

Tekniken för avancerad bildsökning blir allt bättre i takt med att tekniken för att analysera och automatiskt beskriva (eller definiera) innehållet blir allt bättre. Slutligen talade Steve om "spatially-aware retrieval", ett område där det verkligen finns starka drivkrafter för produktutveckling eftersom "85 % av dokumenten på webben innehåller någon form av 'spatial references' ". (Han hade tänkt fortsätta med att beskriva vad personalisering innebar - men det hann han inte ...).

### ***Paneldiskussion***

Efter föredragen fanns lite tid för frågor från publiken under ledning av sessionens moderator Howard Rosenbaum, Indiana Univ., USA (Till vänster på fotot. I övrigt från vänster Steve W, Mikael, John). Den första frågan var: Hur kommer Googles nya Search Appliance att påverka sökindustrin? John svarade att produkten kan komma att få betydelse vad gäller allmänna ändamål men inte för organisationer vars system för sökning är av strategisk betydelse.



Någon i publiken uttryckte skepsis mot automatisk textanalys. Whittaker svarade att för de ändamål han angett så ställs inte krav på exakt eller ens bra analys utan vad det gäller är att identifiera "rätt ställen" i textmassorna för att sedan kunna analysera dessa delar djupare (t ex manuellt). Det är sammanhanget som är det viktiga. I exemplet med kundtjänst pekade han på att AT&T nyttjar en tal-till-texttjänst för automatisk analys av innehållet i röstbrevlådor.

Med anledning av de söksystem som är på gång från IBM och Microsoft ställdes så frågan om de verkligen kommer att bli något att räkna med. Mikael svarade att Microsoft hittills inte tagit sökning på allvar och trodde inte heller att sökning trots de stora satsningar som görs inom projektet Longhorn skulle bli något att räkna med.

John å sin sida menade att det närmaste året blir mycket intressant vad gäller området sökning i och med att Microsoft "bygger in" sökning sina produkter, t ex Windows och MS Sharepoint. Detta kommer att utgöra ett allvarligt hot mot Googles söksystem. Han trodde dock att det tar 5, kanske 7 år innan Microsoft har någon verkligt bra enterprise-sökprodukt ute på marknaden. Whittaker instämde: delar av Longhorn har ju försetats i flera omgångar. Men redan integrationen av sökning i Windows kommer att bli ett lyft eftersom samma sökning kommer att kunna göras i t ex Word som i Outsell.

## ***Ett par svenska föredrag i övrigt***

Som nämnts presenterades tre föredrag av svenskar. Ett har refererats ovan. Trots att jag inte åhörde de andra två följer här, "för rättvisans skull", korta referat, baserad på deras dokumentation.

### ***Managing knowledge through everyday activities***

**Dick Stenmark**, Informatik, Göteborgs universitet, talade om Managing knowledge through everyday activities. Rubriken sammanfattar rätt bra hans erfarenheter. Dicks paper utgår från de studier och experiment han gjort inom Volvo IT. Hans slutsatser vad gäller att skapa och underhålla ett pålitligt KM-system är att man ska tillvarata så mycket som möjligt av det "vardagliga" informationsflöde och informationsanvändning (allt från e-mail och PM till söksystems användning). Ett KM-system ska inte installeras som ett separat system som måste hanteras speciellt utan det ska vara ett integrerat system som automatiskt fylls på med information.

### ***Official business information and the expanding EU challenges & opportunities***

**Håkan Axelsson**, EuroInfoPool/InfoData, talade över ämnet Official business information and the expanding EU challenges & opportunities. Med "expanding" avsågs de nya EU-medlemmarna. De som sökt information online om företag i dessa länder har märkt att det inte är så lätt. Visserligen har de alla statliga företagsregister - men i flera fall går de inte att nå online. Så här är läget: Lettland - Ja, Litauen - Nej, Estland - Ja, Slovenien - Nej, Tjeckien - Ja, Slovakien - Ja, Polen - Nej, Ungern - Ja, Cypern - Ja, Malta - Ja. Ändå ger detta en för positiv bild av verkligheten, ty i flera fall är registren lokalt distribuerade, finns bara sökbart på ett lokalt språk eller så är åtkomsten belagd med restriktioner.

Ett nytt EU-direktiv (2003/58/EC) stipulerar att varje EU-land per den 1/1 2007 ska ha ett centralt, elektroniskt, företagsregister. Det innebär dock inte att allt plötsligt blir bra då, ty direktivet anger varken att en viss standard ska följas eller att registren ska bli tillgängliga externt. Men Håkan hoppades förstås att "hans" portaltjänst EuroInfoPool, på sikt ska kunna ge tillgång till även dessa - utöver de 15 länders företagsregister som redan finns tillgängliga där via samverkansorganet EBR, European Business Register.

## ***Conference Wrap and Close***

Konferensen avslutades med ett ytterst kort och illa förberett tal av en för mig okänd inhoppare (istället för Martin White) inför en mycket fåtalig publik (de flesta hade uppenbarligen insett att vad som komma skulle, dvs inget av värde). Tyvärr en inte så värdig avslutning på detta i övrigt - och det tål att upprepas - mycket välskötta evenemang.

-----

Detta material är i helhet och dess delar skyddat av upphovsrätt. Vidaredistribution, mångfaldigande eller återpublicering kräver skriftligt tillstånd från Svensk förening för informationsspecialister, Box 55580, 102 04 Stockholm. Telefon 08- 678 23 20, Fax 08-678 23 01, E-post kansliet@tls.se