



## Online- och webbnyheter

Nr 2005:6 2005-12-30

### *Rapport från Online Information 2005, London, 29 nov - 1 dec 2005*

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

Gunnel Stjernvall, ISIM, [gunnel.stjernvall@isim.org](mailto:gunnel.stjernvall@isim.org)

*"Business as usual". Så karakteriserar Dick Kaser (Information Today) den 29:e upplagan av Online Information, Online 2005. Och det är bara att hålla med. Inga sensationer, inga braskande nyheter, ... ja - allt var som vanligt. Men så rör det sig ju om ett moget evenemang i en mogen bransch (nu väl inne i sitt 4:e decennium!). Om man ändå ska tala om något nytt så var det de många inslagen om wikis och bloggar på konferensdelen – vars inslag i övrigt i stor utsträckning rörde informationshantering, informationsarkitektur och söksystem, medan "information" (onlinetjänster, databaser, innehåll ...) nästan helt lyste med sin frånvaro. Även det ett tecken på onlinebranschens mogenhet?*

*Utställningsdelen var sig till förväxling lik de närmast föregående årens: lika stor, med nästan samma utställare på nästan samma platser. Det enda man höjde på ögonbrynen åt var att Factiva uteblev men att LexisNexis å andra sidan återvänt efter ett par års frånvaro – och att Dialogs "delning" så tydligt manifesterades (genom att knappt nämnas vid namn och med "delarna" i olika montrar). Javisst ja, en ordentlig skillnad mot förra året märktes trots allt: det rådde ingen "Googlemania" i år - trots att Google deltog med en monter.*

#### **Innehåll:**

<b>1. Allmänt</b>	<b>sid 1</b>
<b>2. Konferensen</b>	<b>sid 4</b>
<b>3. Utställningen</b>	<b>sid 6</b>
<b>4. Utställare</b>	<b>sid 8</b>
<b>5. Föredrag</b>	<b>sid 20</b>

## **1. Allmänt**

Det går knappt att undvika att inleda rapporten från 2005 års Online Information med att återigen skriva att "det mesta var som vanligt"! Något som i och för sig inte är konstigt för detta evenemang, som nu ägde rum för 29:e gången och hela tiden haft samma bransch i fokus; denna bransch som nu nästan halvvägs in i sitt 4:e decennium är i högsta grad mogen.

Huvudevenemanget med konferens och utställning ägde således rum under tre dagar, den 29 november - 1 december, i konferensanläggningen Olympia i Londons västra del. Arrangör var, också som

vanligt, VNU Exhibitions Europe (som tillsammans med Learned Information ägs av VNU Business Media Europe (som också publicerar Information World Review)).

Konferensen, med ca 80 föredrag och paneldiskussioner i tre parallella sessioner, bevistades av ca 720 delegater. Utställningen, som ägde rum i Olympias Grand Hall, hade ca 250 utställare och räknade in 9.800 besökare. Inom ramen för utställningen erbjöds också närmare 100 halvtimmes gratis föredrag. Dagen före huvudevenemanget anordnades 4 heldagars workshops kring ämnena Web Search Update, Content Strategy and Modelling, The Google Legacy och Mining the Deep Web for Company Information, samtliga med ledande branschpersonligheter som föredragshållare.

### ***Online en branschspegel – som vanligt***

Jag brukar tala om Online som en branschspegel. Online 2005 var inget undantag. Som jag nämnde ovan är onlinebranschen i högsta grad mogen. Det sena 1990-talets turbulens, präglad av webben, många nya aktörer (ofta med gratisinformation) och IT-nyan har sedan sekelskiftet gradvis gett vika för alltmer sans och en lugnare marknad. Det händer helt enkelt inte så mycket och åtminstone de största aktörerna på den internationella arenan står stabila (och, vågar jag nog drista mig till att säga, i ett tillstånd av stagnation). Aktörer som Elsevier och Thomson m fl växer bara långsamt. Uppköpseran är slut och nu konsoliderar man.

Utställningen speglade väl denna situation. Visserligen var ca 40 av de 250 utställarna, 16 %, nya sedan förra året (medan lika många försvunnit) - men eftersom bara några få av dessa hyrde stor monterytta så var den ytmässiga förändringen inte alls lika stor. Jag gissar på högst 10 %. Faktum är också att delar av utställningen var sig till förväxling lik förra årets i så motto att många aktörer, i synnerhet de större, hade sin monter placerad på i stort sett samma plats som året innan. (Som kuriosa kan nämnas att "till förväxling lik" tog sig ett konkret uttryck i att utställningsansvarige Victoria Bush i utställningsguiden välkomnade till det "30th Online" och fortsatte med att skriva att "some of you may remember the first Online back in 1976". Lite svårt eftersom det första ägde rum 1977!).



Dock: ser man onlinebranschen med lite bredare glasögon så visar den förstås upp ett helt annat ansikte. Jag syftar här på det som sker vad gäller informationshantering, informationsarkitektur, sökmotorer, teknik (datorer, verktyg, kommunikation etc - IT kort och gott), Open Access, företeelser som bloggar och wikis, aktörer som Google, Yahoo! och MSN, mm. Men allt detta har ännu inte förmått ändra den "centrala" branschstrukturen – där det bärande fortfarande är "information". Detta speglades också på utställningsdelen av Online, där informationsaktörer (onlinetjänster, databaser, rapportproducenter, förlag etc) dominerade men där det självklart också fanns utställare inom de andra områdena.

En intressant iakttagelse vad gäller onlinebranschen, vilket också avspeglades på Online 2005, är att det tycks vara ont om information ... Det tar sig uttryck i att samma information finns tillgänglig (via t ex licenser eller återförsäljaravtal) hos ett flertal aktörer, det må sedan röra sig om primärinformation (t ex tidskriftsartiklar, marknadsrapporter), referenser (t ex i databaser), monitorerad information från webben, företagsinformation eller hela databaser. Det här blir så uppenbart när man som på Online har utställarna samlade på ett ställe. Företeelsen i sig är såklart inte ny - t ex har databaser som Medline, Compendex etc alltid funnits vid många onlinetjänster - men den tycks ha accelererat avsevärt på senare tid.

Att det är "ont om information" tar sig också uttryck i att många onlinetjänster satsar på att krama ett så stort "värde" som möjligt ur den information man redan har. Ett traditionellt sätt är att förpacka den på olika sätt. Det är precis det som Elsevier gör med sin Scopus, som ju till mycket stor del består av referenser i andra databaser, och som LexisNexis gör med sina nya produkter för "reputation management", dvs återanvänder sådan monitorerad information som redan finns tillgänglig i andra tjänster.

Ett annat sätt är att göra informationen till föremål för intelligent återanvändning - vilket visserligen också är traditionellt men som nu verkligen blivit högsta mode. Jag syftar här på analyser; verktyg för analyser av funnen information, ofta i kopplad till grafiska redovisningar och annan visualisering. Exempel är patentanalyser, medianalyser (t ex av det egna företagets "synlighet", jämförelser med konkurrenter etc), företagsjämförelser, branschens utveckling, makroekonomiska analyser, marknader, citeringsanalyser, kemiska formler/strukturer ... Faktum är att just analysvertyg av olika slag hade en mer framträdande roll än "information" hos många av utställarna.

### **Mer information om Online 2005**

Vill du ha ytterligare information om Online 2005 kan du nå den "officiella" informationen om evenemanget (program, utställare, pressinformation, seminarieinformation, utnämningar etc) via [www.online-information.co.uk](http://www.online-information.co.uk). Här kan också nämnas en nyhet för årets Online, nämligen den konferensblogg som arrangören introducerade någon månad innan evenemanget. Den gav möjlighet för sessionsledare, föredragshållare och andra med någon funktion på konferensen att ge synpunkter, diskutera, kommentera egna och andras föredrag, etc. Resultatet blev dock mycket magert, även om man säkert kan få ut ett och annat via de ljudklipp som finns med där. Se [blog.online-information.co.uk](http://blog.online-information.co.uk).

Betydligt mer innehållsrik är Information Todays blogg på [www.infodayblog.com](http://www.infodayblog.com), som nu alltså återkom för tredje året. Här finns inlägg av onlinekändisar som Marydee Ojala, Don Hawkins, Nancy Garman, Dick Kaser och Paula Hane. Den innehåller förutom mängder av kommentarer, referat och intressanta reflektioner ett flertal foton från utställningen och kan verkligen rekommenderas!

Slutligen kan här nämnas att FreePint nr 196 innehåller personliga reflektioner från några deltagare, varav en förstagsbesökare. Se [www.freepint.com/issues/151205.htm#feature](http://www.freepint.com/issues/151205.htm#feature).

### **Online 1999-2005 - lite statistik**

Det är nu sjunde gången Online- och webbnyheter utgörs av en rapport från Online. Här nedan presenteras en sammanställning av några uppgifter från dessa rapporter.

	<b>Konferensen</b>					<b>Utställningen</b>		
	<b>Antal delegater</b>			<b>Antal föredrag *)</b>		<b>Antal utställare</b>		<b>Antal 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 2002</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>
<b>Online 2005</b>	<b>720</b>	<b>23</b>	<b>29</b>	<b>90</b>	<b>1</b>	<b>250</b>	<b>1</b>	<b>9.800</b>

< Något under

> Något mer än

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

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

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

\*\*\*\*\*) Uppskattning

/Lars

## 2. Konferensen

Konferensen bestod av ca 70 föredrag (varav 10 keynotes) och 6 paneldiskussioner i tre parallella spår. Spåren var indelade i sessioner om 1.5 timmar och bestod i normalfallet av tre föredrag, ibland med en inledande keynote. Under lunchtid förekom också s k roundtables, totalt 6 st med följande teman: Competitive Intelligence, eLearning, Taxonomies, Wikis and blogs, Selling your brand och Communities of practice. Därtill ägde 6 s k ”international forums” rum: German, Italian, Spanish, French, Dutch och Nordic (där den sistnämnda handlade om “Scholarly Scommunication – future and challenges” och leddes av Gunilla Widén Wulff, Åbo Akademi).

De områden som utmärkte sig genom flest föredrag var

- informationshantering och -arkitektur
- sökmotorer (desktop, enterprise)
- wikis, bloggar och RSS

Ska man tala om ett tema för konferensen så var det just wikis och bloggar – och det var också det område som syntes vara mest populärt bland delegaterna. Vid något tillfälle var det till och med så många som ville delta att lokalen inte rymde alla som ville in.



Naturligtvis fanns det föredrag inom en rad andra områden, t ex portaler, bibliotek, Kina, upphovsrättsfrågor, publicering, Open access - där det kan noteras att det sistnämnda var mer nedtonat än de närmast föregående åren. Mest notabelt var dock den nästan totala frånvaron av föredrag och andra inslag som behandlade ”information”, i alla fall kommersiell sådan (innehåll, onlinetjänster, databaser, marknaden för ...). Det närmaste var en paneldiskussion som behandlade sökning. Ja, det är inte utan skäl som arrangörerna numera kallar Online Information för ”The world’s no. 1 event for content and information management solutions” ...

### *UK och USA i dominans*

Koncentrationen till UK och USA vad gäller föredragshållarnas hemvist var ännu mer markant än tidigare år. Av de 80 föredragshållarna med sina uppgifter i Proceedings var 37 från UK, 23 från USA och 20 från andra länder. Två var från Danmark, Frankrike, Schweiz, Tyskland och Nederländerna och en från Norge, Sverige, Österrike, Belgien, Luxemburg, Slovenien, Libanon, Kina, Sydafrika och Australien. I realiteten deltog ytterligare ett tiotal, t ex som keynote-föreläsare eller medförfattare (bl a en finländare). Det enda svenska inslaget utgjordes av mitt eget föredrag “Troubles and joys - the evaluation, selection and implementation of an enterprise search engine for a large online information service”, som jag återkommer till nedan.

Konferensen hade något fler delegater än förra året; omkring 720. De kom från närmare 40 nationer. Störst var givetvis Storbritannien med 297, vilket var 10 % mer än 2004. Därefter kom USA med 94, vilket var mer än en fördubbling. Därpå följde Nederländerna med 37, Norge med 32, Finland 29, Frankrike 25, Sverige 23, Tyskland 21, Danmark 20, Schweiz 15, Sydafrika 12, Österrike 11, Spanien 10, Belgien 9, Ungern 8, Kanada 8, Irland 8, Italien 7, Slovenien 6, Island 6, Australien 5, Israel 5 och Japan 5. Ingen deltog från Ryssland, vilket nog är första gången. Europa hade totalt 576 delegater (varav Norden 110), Nordamerika 102, Asien 25 och Afrika 17. Sydamerika var inte alls representerad. Precis som beträffande föredragshållare har koncentrationen till UK och USA ökat.

Nästan två av tre av oss svenskar och finländare var från offentlig sektor, mest universitet. Av 23 svenskar är 14 medlemmar i Svensk Förening för Informationsspecialister, SFIS.



Konferensavgiften var något lägre i år än förra året, nämligen som lägst 8.100 kr vid bokning före 16 september och som högst 10.500 kr vid anmälan efter den 2 november. Delegater från offentlig sektor



slapp undan med 5.200 kr hur sent de än anmälde sig. Prissänkningen skedde dock till priset (!) av att konferensavgiften inte inkluderade konferensmiddagen, den s k "Awards Ceremony", vilken den alltid tidigare gjort. Eftersom en middagsbiljett kostade 1.700 kr så hade priset i realiteten höjts - det ordinarie lägsta priset vid Online 2004 var nämligen 9.000 kr.

#### *Konferenslobbyn*

Trots priset (OK, ett antal av dem var förstås bjudna) deltog närmare 400 personer i onsdagskvällens Awards Ceremony, vilket var i stort sett lika många som tidigare år. Förutom fin middag vankades här musik och dans – och så förstås den traditionsenliga utdelningen av "awards"; i år fler än någonsin: 18 st. Det rörde sig om allt från "Innovation in knowledge management" och "Best STM information product" till "Best implementation of a business blog". Bland utmärkelserna vill jag speciellt nämna "Online Information Lifetime achievement award", sponsrad av Business Intelligence Services (dvs "halva" Dialog), eftersom den i år utdelades till Martin White, bland annat för hans idoga arbete som ordförande för Online-konferenserna fr o m år 2000. Martin, som deltagit i nästan alla Online sedan starten 1977, har för övrigt meddelat att Online 2006 blir hans sista som ordförande.

#### *Tunn dokumentation*

Dokumentationen var tyvärr lika "tunn" som förra året. Proceedings innehöll bara ca 40 presentationer, varav högst 30 kan betraktas som något så när fullödiga "papers". De resterande utgjordes av enbart abstracts eller, i bästa fall, de bilder som visades. Det innebär ungefär hälften av samtliga "möjliga" föredrag. Det här upplevde jag som mycket märkligt, i synnerhet mot bakgrund av min egen erfarenhet som föredragshållare. Förhållningsreglerna var nämligen mycket strikta, med inlämning av manus senast i augusti och de bilder man skulle visa senast tre veckor innan konferensen.

Proceedings kan köpas för 1.600 kr. I vilken mån dokumentation också i år kommer att publiceras på konferenswebben vet jag inte. I så fall kan möjligen en del av det saknade materialet dyka upp där.

/Lars

#### **ICLG medlemskväll**

ICLG höll en medlemskväll under konferensens första dag. ICLG står för the Industrial & Commercial Libraries Group och är Storbritanniens CILIP-anslutna motsvarighet till SFIS. (CILIP = The Chartered Institute of Library and Information Professionals). Anledningen till kvällens mingel var lanseringen av deras nya medlemsbrev, som inte helt överraskande heter ICLG news.

För £ 25.00 kan man bli medlem i enbart ICLG, utan att gå via ett CILIP-medlemskap. Se [iclg@cilip.org.uk](mailto:iclg@cilip.org.uk) eller [www.iclg.org.uk](http://www.iclg.org.uk). Eftersom detta är SFIS närmaste motsvarighet i Storbritannien vill jag glädjande förmedla att deras ordförande, Oriole Newgass, välkomnar framtida samarbete i någon form.

/Gunnel

### 3. Utställningen

Som nämnts var antalet utställare ca 250, dvs ingen förändring jämfört med de senaste åren. Av dessa var 40 nya för i år (dvs drygt 15 %), vilket också var som vanligt. Flertalet av dessa 40 hade bara små montrar vilket innebar att den ytmässiga förändringen var mindre (ca 10 %). Detta i kombination med att många av förra årets utställare, i synnerhet de större, hade sina montrar placerade på i stort sett samma plats som förra året gjorde utställningen i många stycken sig till förväxling lik utställningen vid Online 2004.

Bland de nya utställarna märktes kanske mest LexisNexis Butterworths och Google; LexisNexis eftersom de återkommit efter två års frånvaro och Google därför att ..., tja eftersom det var Google (som dock deltog med en mycket liten monter). Bland andra nya fanns Vivisimo, Microsoft, Infocandy, Canon, EIBS och MediaSurface. D&B var också med efter ett par års frånvaro – men nu tillsammans med Hoover's. En svensk aktör deltog, nämligen mitt eget företag Infodata som visade EIP, EuroInfoPool, dvs portalen för åtkomst till bl a företagsinformation från officiella källor i Europa.

Mest notabelt bland de som saknades bland dem som varit med tidigare var Factiva. Jag var inte den ende som efter att ha letat efter men inte funnit dem på utställningsgolvet knappt trodde våra ögon när vi inte heller fann dem i utställarförteckningen. Bland de som saknades av de som varit med tidigare (men inte förra året) fanns Autonomy, Verity, HW Wilson, Skyminder, Bloomberg och ICC. Vad gäller sökmotorer upplevde jag att de var färre än tidigare. Utöver de nämnda uteblev bl a Mondosoft, som deltog förra året. Inslagen av nationella onlinetjänster (dvs sådana som primärt är inriktade på sin hemmamarknad) var i princip obefintligt - bortsett förstås från UK. Som exempel kan nämnas att tyska aktörer som tidigare varit med, som GBI och Genios (numera i kartell) och Dimdi, fortsätter att utebli.



Faktum är att utställarna i större utsträckning än tidigare representerade området STM och hade en inriktning mot den akademiska sektorn. En trend i denna riktning har ju följt pågått under de senaste fem åren men jag upplever den nu mer markerad. Området finansiell information försvann för flera år sedan och rättsinformation var nästan inte alls representerat.

Information Today publicerade nyligen sin nya EContent 100-lista över "informationsindustrins mest betydelsefulla företag". Av de hundra aktörerna var 28 representerade på Online 2005, nämligen Alacra, Alfresco, Chemical Abstracts Service (CAS), CSA, D&B, Day

Software, EBSCO, EMC Software, eMeta, Endeca, Fast Search & Transfer, Google, Ingenta, Knovel, MarkLogic, Mediasurface, Microsoft, Moreover Technologies, OCLC, Ovid Technologies, ProQuest, Questel-Orbit, Reed Elsevier, Springer, Swets, Thomson, Vivisimo och Xrefer. Det är i exakt nivå med förra året. Kanske kan man på basis av detta säga att Online representerar drygt en fjärdedel av industrin – vilket definitivt är mer än några andra evenemang, inklusive de amerikanska.

För två år sedan skapades "utställningen i utställningen" CME (Content Management Europe) och förra året EDRM (Enterprise Document & Records Management Europe) som del av Online. Dessa återkom i år och var förlagda i borte delen av utställningshallen. 42 av utställarna hade valt att exponeras under dessa rubriker. I själva verket var denna åtskillnad inget man som besökare lade märke till. Många utställare i den "ordinarie" utställningen kunde lika gärna ha funnits i CME/EDRM-delen.

Förhållandet mellan utställare med "information" och "övrigt" (dvs primärt "icke-information") kan grovt sett uppskattas till 3:2. En sammanställning av utställarnas enligt deras egen kategorisering (de fick välja en eller flera kategorier) ser ut så här:

<i>Kategori</i>	<i>Antal utställare som uppgav kategorin</i>
<b>Company Business &amp; Financial Information</b>	<b>41</b>
<b>Legal &amp; Regulatory Information</b>	<b>22</b>
<b>Market Research &amp; Competitive Intelligence</b>	<b>34</b>
<b>Recruitment &amp; Training</b>	<b>10</b>
<b>Scientific, Technical &amp; Medical</b>	<b>83</b>
<b>Content Management</b>	<b>95</b>
<b>Enterprise Search Solutions</b>	<b>36</b>
<b>ePublishing Services &amp; Solutions</b>	<b>76</b>
<b>e-Records Management &amp; Document Management</b>	<b>44</b>
<b>Intranets &amp; Portals</b>	<b>60</b>
<b>Library Management Solutions</b>	<b>40</b>

Ytmässigt var dock utställare med "information" i stor övervikt. Bland dessa fanns ju giganter som Wolters Kluwer, EU-kommissionen, Thomson och Springer. Bland övriga fanns i alla fall en med stor monter, nämligen Innodata Isogen, ett företag som arbetar med digitalisering, datakonvertering mm och med verksamheten till stor del förlagd till Asien.

Inom ramen för utställningen erbjöds också närmare 100 st halvtimmes gratisföredrag, varav många givetvis utgjordes av till föredrag förklädda produktpresentationer men där många också var av mer allmän karaktär. STM-spåret innehöll t ex föredrag som STM information industry outlook, What's new in medical resources och Update on engineering information resources (hölls av Roddy MacLeod). Business Information-spåret innehöll bl a föredragen Top 10 online trends (hölls av Marydee Ojala), How to locate statistical data successfully, Business information in China och Successful contract negotiation. Teoretiskt sett kunde den som så önskade fylla sina tre dagars besök på Online 2005 med enbart gratisföredrag. I praktiken var detta dock knappast möjligt, dels pga "lokalernas" tillfälliga karaktär, dels pga att dessa lokaler i vissa fall snabbt blev fullsatta.

Det sistnämnda blev jag själv drabbad av. Med anledning av det minst sagt magra inslaget av "information" på konferensen siktade jag in mig på paneldiskussionen "Business Information Under the Spotlight – Future predictions", med representanter för D&B, LexisNexis och Bureau van Dijk i panelen. Men se där bedrog jag mig - det var fullsatt.

/Lars

## 4. Utställare

Har inget hänt på 20 år? Samla in, paketera och leverera. Är det inte uttryck vi hört i många år? Är det bara kläderna som är nya; paketeringen alltså. Samtidigt som detta känns lite segt och gammalt träffar jag kollegor i pressrummet som förvånat utbrister: - bibliotek ... ja, jag kände en gång en person som jobbade på bibliotek.

Men vi har ju XML. Eller som en kollega uttryckte det apropå Gartners årliga rapport om tekniska nyheter: de flesta av dessa tekniska nyheter brukar göra sitt intåg på biblioteken förr eller senare.

Av de 250 utställarna på Online 2005 uppgav 86 st i utställningsguiden att de skulle presentera en "New Product". Måhända borde de aktuella företagens monteransvariga också ha berättat detta för sin personal innan öppnandet ...

Mer information om samtliga utställare inkl länkar till deras webbplatser återfinns via den kompletta utställarförteckningen på [www.online-information.co.uk/cgi-events/exhibitors.pl?exhibition\\_id=79](http://www.online-information.co.uk/cgi-events/exhibitors.pl?exhibition_id=79)

I det nedanstående redovisas intryck från ett 40-tal utställare. Uppdelningen är löst baserad på utställningsguiden och utställarna presenteras under den rubrik de själva använt för presentationen av sitt företag. När ni hittar IEEE bland ePublishing-företagen är det alltså deras egen presentation detta baserar sig på.

/Gunnel

- <b>Content Management, Information Management</b>	<b>sid 8</b>
- <b>Market Research, Business Intelligence</b>	<b>sid 9</b>
- <b>e-Publishing</b>	<b>sid 13</b>
- <b>Scientific, Technical &amp; Medical</b>	<b>sid 15</b>

## ***Content Management, Information Management***

### ***Emantras, Inc.***

Det känns välkomnande att börja med ett företag som Emantras. Deras monter var en väl fungerande av det mindre slaget, med väl genomtänkt innehåll och trevligt bemötande. Emantras är ett utvecklingsföretag inom Media och Content, fokuserar på verktyg för e-learning (mycket medicin) men även företagsutbildningar. Man utvecklar tekniken i Indien, men företaget finns även i Kalifornien. Bland de senaste kunderna räknar man NASA, Siemens, HP, Blackwell Publishing, Elsevier, McGrawHill, Danfoss, m fl.

/Gunnel

### ***EOS International***

EOS International har för tredje året i rad hamnat på en lista över de 100 snabbast växande privatägda företagen i och runt San Diego. De har sålt hittills till fler än 3000 Special Libraries, och har under året även annonserat att EOS system redan är förberett för ett 13-siffrigt ISBN. Deras EOS.web finns f o m i år även i fransk version. Man vill gärna förmedla till svenska informationsspecialister att man under året etablerat en state-of-the-art EOS Global Data Center vid deras huvudkontor i Carlsbad, Kalifornien. Här finner vi library automation solutions för bibliotek över hela världen. Deras kunder finns t ex i Irland, UK, Switzerland, Israel, Australien och förstås Nordamerika och Kanada.

/Gunnel

### ***Fast Search & Transfer***

Norska FAST kommer nu med tekniska lösningar för att kunna göra sökningar från mobiltelefonen. Man har också lanserat produkten FAST ProPublish, som gör det möjligt att från en plattform, samla in, organisera, leverera information.

FASTs CEO, John M Lervik har under året hamnat på Forbes lista över World Technology Innovators. FASTforward 06 är en konferens FAST anordnar i Miami, Florida den 1-3 februari. Enligt Lervik (dn.no - nov. 05) är FAST ett av världens idag snabbast växande företag. Företaget kommer också på tredje plats på norrmännens aktietips inför 2006. FAST finns på OSE, Norges börs, sedan 2001.

/Gunnel

### ***Immediacy Content Management***

Immediacy är ledande i UK när det gäller CMS (Content Management Software). Företaget har under året lanserat ett nytt DMS (Document Management System), som svar på att kunderna efterfrågar enklare och billigare system. Detta har bl a Gartner Group, och andra liknande källor, rapporterat som en trend; att företagen nu efterfrågar enklare system. John Greenless, arkitekten bakom produkten, säger att den erbjuder ett enklare alternativ för organisationer som vill hantera alla typer av dokument. Bland kunderna finner vi bl a BBC, British Council, Avis, Hilton Hotel Group m fl.

/Gunnel



### ***KnowledgeView Ltd***

Man levererar sedan ett tiotal år till Independent News and Media Group (INM) ett system för såväl författande som arkivering: Editorial Sharing and rich-media Archiving (ESA). Med RAPID browser omvandlar man pdf till XML. Nu integrerar man RAPID browser med Adobe Incopy och InDesign och får RAPIDPublish.info. Under the Guardian Newsroom in London den 29 mars 2006 lanserar man RAPIDPublish.info tillsammans med Adobe. KnowledgeView hjälper mestadels utgivare som BBC och CNN men även Marks & Spencer och UNICEF.

/Gunnel

### ***Solcara Limited***

Solcara, som öppnade London-kontor under 2005, har sin huvudsakliga tekniklösning i Solcara Information Manager (SIM). Kortfattat består systemet av en mängd moduler som tillsammans håller ordning på en organisations tillgångar inom information, innehåll och kunskap. Denna tjänst inkluderar även deras Solsearch, som innebär att man söker i samtlig information på en gång. Solcara har sedan hösten 2005 120 kunder, av vilka 20 finns i Europa utanför UK. Man har helt nyligen öppnat försäljningskontor i Köpenhamn, och säljer i bl a Sverige. I början av 2006 har man ambitionen att ta sig in på den amerikanska marknaden.

/Gunnel

### ***TFPL Ltd***

TFPLs monter var stor och välplacerad, vilket sänder ut en massa positiva signaler. Deras press-folder var desto tunnare. Inte mycket av händelser att rapportera egentligen. Måne detta bero på den relativt nya ägaren; IDOX. Den som lever får se. Bland nyheterna från 2005 kan vi dock saxa en ökande rekryteringssatsning inom kontrakt-jobb (projektanställningar), nytt kontor i Manchester och att man håller sin EBIC konferens i Milano i mars 2006.

/Gunnel

## ***Market Research, Business Intelligence***

### ***Alacra, Inc.***

I Alacras monter fick jag information och broschyr. Det är en intressant vandring på Online-golvet, en del ger ett bra mottagande eftersom de inser att som Press kommer man att förmedla intrycken (och ibland även vad de säger) till läsarna. Och ytterligare en del tycks inte förstå det sambandet alls. Bland nyheterna framhåller man sin Alacra Alert, där man kan få nyheter per e-post varje dag eller varje vecka, från Alacras källor.

I september i år lanserades Alacra Wiki på [www.alacrawiki.com](http://www.alacrawiki.com), en gratis webbplats som hänvisar till resurser inom "business information" i form av gratis och avgiftsbelagda webbplatser, databaser, onlinetjänster, directories, nyheter mm. Informationen kan uppdateras av allt från leverantörer till användare.



Alacras onlinetjänst (Alacra.com) ger tillgång till en imponerande samling databaser med företagsinformation, drygt 100 st, vilka är upplagda vid ett eget system i New York. De kommer från D&B, Thomson Gale, Thomson Financial, ICC, Reuters, Nikkei, BvD, EIU, Mintel, S&P etc. Men smakar det så kostar det: abonnemang kan kosta från över 100.000 kr per månad och uppåt. Man har några kunder i Sverige.

/Gunnel, Lars

### ***Bureau van Dijk Electronic Publishing***

Det är alltid lika trevligt och värdefullt att komma till en monter – och ett företag – som förmedlar så tydligt att de är professionella på alla sätt. Stor modern monter, bra personal, som tar sig tid att prata och presentera det nya, ytterligare personal som tar sig tid att svara på mina uppföljande mail. Helt enkelt ett företag man skulle vilja göra affärer med.

Konkurrentbevakarverktyget MINT har under året fått tillökning i form av Orbis-databasen med sina 15 miljoner företag, och detta kallar man MINT Global. MINT Global kommer att lanseras i början av 2006 och förhandsvisades under Online konferensen.

Ytterligare tillökning i MINT-stallet blir under 2006 MINT Spain och en databas innehållande ryska och ukrainska företag; Ruslana. Och med detta visar det sig att de är inte enbart en mintgrön utställare, utan också har gediget och innovativt innehåll.

/Gunnel

### ***Business Insight***

Här kom man till en liten monter, med en (1) person i. Business Insight är ett litet företag, med kunder i Storbritannien. Affärsidén går i princip ut på att vara företagsbibliotek åt alla som blir så kallad prenumererande kund. En himla bra idé. Kanske något för oss i Sverige att ta efter. UKs eget informationssökeri ligger i Birmingham och hjälper mestadels lokala företag i deras behov av information.

/Gunnel

### ***Business Intelligence Services***

Business Intelligence Services? En ny aktör? Ja, på sätt och vis – men under detta namn på Online 2005 “dolde” sig faktiskt Thomson Business Intelligence, TBI, som i sin tur är det (ganska) nya namnet för den del av Dialog som består av NewsEdge, Intelliscope och Profound etc. Business Intelligence Services hade en egen monter skild från de andra Thomson-montrarna på Online, Thomson Gale och Thomson Scientific. Lite märkligt kändes det ...

Nåväl, TBI säger sig erbjuda en smart ”single solution” för finansiell information, marknads- och nyhetsinformation. Man vill helt enkelt erbjuda en ny snabb tjänst som ständigt ger kunderna den information man behöver för att kunna fatta rätt och bra beslut. Som de uttrycker det: “It’s all about what you know and when you know it”.

/Lars, Gunnel

### ***Business Monitor International***

Business Monitor visade gällar återigen sin ”nyhet” Daily Alert Service inom finansiell- och marknadsinformation, med levererans via e-post.

/Gunnel



### ***Butler Group***

Butler är numera ett Datamonitor-företag. Butler erbjuder forskning, analys och råd inom IT-området och har så gjort i bl a Computergram-form i 20 år. För den som vill hålla sig ajour inom IT-utvecklingen är det nog ingen dum idé att läsa Butler Group Review regelbundet. I mars håller Butler ett Master Class-seminarium i Stockholm och kommer att tala om Achieving IT flexibility.

/Gunnel

### ***Domex e-Data Pvt Ltd***

Domex ligger i södra Indien och sköter editering, abstract, indexering m m åt t ex många av de databasföretag som finns representerade på Online. Ett sådant som Dir. Balakrishnan nämnde är Inspec. Domex är det andra Indien-baserade företaget jag träffar på under Online, vilket är ett tecken i tiden och på att det säkert finns fler.

/Gunnel

### ***EuroInfoPool***

Svenska EuroInfoPool ställde ut på Online för tredje gången. Man erbjuder samordnad sökning i och åtkomst till information i de officiella företagsregistren i 15 länder i Europa. Det rör sig om allt från svenska Bolagsverkets Näringslivsregister till brittiska Companies House. Åtkomsten till registren sker via direktåtkomst (gateways) vilket innebär att det rör sig om helt färsk information. Totalt kan uppgifter om 18 miljoner företag nås på detta sätt.



Därtill erbjuds kreditinformation om 25 miljoner företag i 19 länder. Nytt för mig var att denna information numera tillhandahålls av fler leverantörer än enbart Graydon. De är Coface (för Bulgarien, Frankrike, Kroatien, Polen, Portugal, Rumänien, Slovakien, Tjeckien, Ungern), Kobmandstandens Oplysningsbureau (Danmark), Burgel Wirtschaftsinformation (Tyskland), Icap (Grekland), Experian (Irland), Eurocredit (Italien), Informa (Spanien) och Kredtiscutzverband Von (Österrike). Graydon fortsätter att vara leverantör vad gäller Belgien, Nederländerna och UK.

EuroInfoPool, som drivs lite i skuggan av giganter som Hoover's, D&B och Skyminder, uppger att man nu drygt tre år efter starten upplever en snabbt ökad efterfrågan på sina tjänster.

/Lars

### ***Euromonitor International***

Under Online 2005 lanserar Euromonitor sin Country Insight. En, som det sägs, lättanvänd internet-tjänst för fakta och analyser om ekonomiska, politiska, sociala och demografiska trender i mer än 70 länder. Sedan tidigare har man även tjänsten The Global Market Information Database som är ett prisbelönt, internetbaserat referensverktyg för information om länder, marknader, kunder och livsstilar.

/Gunnel

### ***Global Insight***

Återigen ett företag som känns helt rätt att träffa. Bra monter, bra personal som träffar en, ger en bra genomgång av deras produkt och nyheter.

Global Insight är ett high profile-företag. De ger, för en inte oansenlig summa, heltäckande länderbevakning innehållande företag, regioner, industrier och marknader. Man har rätt bra på fötterna när man t ex publicerar 10 ekonomiska förutsägelser inför 2006, samt publicerar facit för dito 2005.

Man har nästan 4.000 kunder från industrin, finansvärlden och offentlig sektor. Global Insight finns i 12 länder, med fler än 600 anställda. Det här företaget kändes helt klart intressant för en kommande träff inom SFIS KM-grupp.

/Gunnel

### ***Hoover's, D&B***

Företagsinformationstjänsten Hoover's, som för första gången på Online delade monter med sin ägare D&B (som dock egentligen inte märktes), presenterade inget nytt på Online 2005. Men det kanske inte behövdes; det tycks gå bra ändå för Hoover's med det stora utbud av företagsinformation och de verktyg som man redan erbjuder från egna källor och från D&Bs stora utbud. Att det är så pekar inte minst den uppgift jag fick om att man nu har 500 anställda i USA på. Vid tidpunkten för D&Bs förvärv av Hoover's, dvs vintern 2002/2003, var antalet 230. 10 personer finns vid kontoret i UK.

Den prisstruktur Hoover's införde för ett par år sedan bibehålls. De heter Pro, Pro Plus och Pro Premium, där den sistnämnda är den mest avancerade och som bl a ger tillgång till verktyget Family Tree

Viewer för att identifiera ägarförhållanden. 3.9 miljoner företag täcks av 500.000 sådana "family trees". Men det tycks också som om det går att välja delar ut utbudet. Gratisversionen av Hoover's finns såklart fortfarande kvar – och så erbjuds billighetsversionen Hoover's Lite för studerande och privatpersoner, vilken ger tillgång till uppgifter om 40.000 företag. Som komplement till den rena företagsinformationen erbjuds Pro-kunderna sedan i somras också nyheter om funna företag från 3.000 internationella nyhetskällor.

/Lars

### ***LexisNexis Butterworths***

Efter två års frånvaro var LexisNexis åter på Online, denna gång representerad av dess brittiska representant LexisNexis Butterworths. De slog på stora trumman för den kommande nya "LexisNexis Business and News Online Service". Det är en tjänst som tagits fram som en del i ett globalt utvecklingsprogram med en budget på 1.2 miljarder kr. Så långt jag fann baserar det sig inte på nytt innehåll utan på samma 34.000 källor som tidigare – men tjänsten ger till skillnad mot tidigare möjlighet till sökning i all information samtidigt; något som inte är möjligt i befintliga tjänster. Detta gäller oberoende av om det rör sökning av företagsinformation, marknadsrapporter, litteratur, nyheter etc. Den tidigare begränsningen i antal träffar tas bort - och precis som hos flertalet konkurrenter så införs automatisk kategorisering av sökresultat, KWIC och mycket mer.

Innehållsmässigt riktade LexisNexis Butterworths in sig på de nya tjänster man har för analyser – t ex av hur media bevakar det egna företaget, andra företag (läs konkurrenter), personer, produkter, marknader, varumärken, egna texter (!) etc. Bland dessa tjänster finns LexisNexis MarketImpact, LexisNexis Know Your Customer (KYC) (avsedd för bl a "riskbevakning") och LexisNexis CopyGuard (en tjänst i stil med PrioInfos tjänst "Urkund", dvs upptäckt av textplagiat, men inriktad på att identifiera upphovsrättsintrång). Men även här rör det sig alltså inte om nya källor utan om att nyttja dessa via nya förpackningar och med nya verktyg.

LexisNexis tillhandahåller också en tjänst som kallas "Negative News". Den ger användarna möjlighet att få fram artiklar mm om personer eller företag som är inblandade i olika typer av "skumraskaffärer". Det sker genom att sökningarna automatiskt avgränsas till material där personer/organisationer nämns i samband med konkurser, brottsliga handlingar etc. Kostnaden för denna typ av "investigations" uppgavs vara 700 kr, dock krävs köp av minst 25 sådana per år. Det går också att få motsvarande som bevakning.

/Lars



### ***Moreover***

Moreover, en av de mest kända vad gäller monitorering av webbkällor, var med på Online för vad jag tror andra gången. Man bevakar f n 12.000 källor på webben (vilket genererar 200.000 artiklar per dag), 1.3 miljoner bloggar (vilket ger 80.000 poster per dag) och 18 nyhetskanaler typ BBC och CNN. Det är intressant att jämföra detta med läget vid Online 2004 då man bevakade 9.000 källor och 750.000 bloggar – för att inte tala om att man i slutet av år 2003 bara bevakade 3.000 källor. Och mer blir det: jag fick mig berättat att man under kvartal 1 kommer att vara uppe i 24.000 källor! Även svenska källor ingår: ca 60 st. Apropå Sverige har Moreover samarbete med såväl Magenta (som nyttjar material från Moreover) och Retriever (om vad fick jag inte klart för mig).

Moreover levererar information till fyra typer av kunder: Företag (intern användning), medieaktörer/förlag, webbportaler och enskilda användare. Kunderna får tillgång till ett arkiv bestående av de 3 senaste månadernas material – och de får också om de vill indexera och spara allt material som levererats till dem. Bland kunderna finns MSN, AskJeeves, BBC och BvD. Några svenska kunder har man



också, vad jag förstår större företag som bevakar sina konkurrenter. Moreover, som startades 1998, köptes i höstas av VeriSign. Man har 40 anställda i USA och UK.

/Lars

### ***Scope eKnowledge center***

Återigen ett företag vars kunder i stor utsträckning fanns representerade just här på Online. De tillhandahåller tjänster inom forskning och e-content till stora organisationer och utgivare. Deras stora nyhet under Online är den kioskvältande nya patenterade taxonomien inom kemi, vilken sägs ha mottagits med stor entusiasm inom den kemiska industrin och bland informationsspecialister inom kemi.

/Gunnel

### ***Snapdata International***

Precis som vanligt deltog ett antal primärproducenter av marknads-, bransch- och företagsrapporter på Online, bland dem Business Monitor International (se ovan), Economist Intelligent Unit, Euromonitor International (se ovan), Global Insight (se ovan), Hemscott, Infionals, ISI Emerging Markets, Key Note Ltd, Snapdata, The Freedonia Group och The Wall Street Journal Europe. Här några intryck från Snapdata.

Snapdata International, som tidigare hette Snapshots International (man bytte namn i somras), UK, är bekant för sina marknadsrapporter; detta inte minst pga att de marknadsförs av flera välkända "aggregators", typ Alacra. Kärnan i utbudet är fortfarande just marknadsrapporterna, numera närmare 10.000 st, vilka täcker en rad branscher och regioner – från den brasilianska kaffeindustrin till telekommunikationer i Kina. De hålls kontinuerligt uppdaterade av Snapdatas egen personal, ca 30 personer samt personer med regional förankring på intressanta marknader. Utbudet ökar årligen med omkring 2.000 rapporter. De tillkommer ofta som en följd av speciella beställningar från kunder.

På Online passade man på att lansera en helt ny serie rapporter, nämligen Company Portraits Series. Det rör sig om företagsrapporter som istället för "traditionellt" innehåll av typen finansiella och andra data erbjuder information sett ur "marknadssynpunkt". Det åstadkoms genom att basera dem på material från uppgifter man redan samlat för sina marknadsrapporter (dvs en slags återanvändning). 2004 åstadkom man något liknande när man skapade sin Global Panorama Series för globala och regionala rapporter, f n ca 250 st.

Rapporterna kan köpas direkt via Snapdatas webbplats för ca 2.000 kr/st – om man inte föredrar att nyttja dem via någon återförsäljare. Det går också att teckna fastprislicens för obegränsad tillgång till rapporter under ett år. För hela samlingen gäller 250.000 kr men om man nöjer sig med enbart sådana som är relaterade till Europa 140.000 kr eller enbart Skandinavien för 28.000 kr per år. Det tycks gå bra för Snapdata. Ett tecken på det är att man i somras öppnade ett kontor i New York. Samtidigt menade man att en del av deras konkurrenter har det lite kärvt.

/Lars

## ***ePublishing***

### ***DeltaXML Ltd***

DeltaXML arbetar med förändringar inom XML-miljö. Allt eftersom XML blir mer allmänt känt och använt, från förlag och flygindustrin till övriga segment i samhället, har behovet av att kunna kontrollera förändringar i XML vuxit högst avsevärt enl DeltaXML. Företaget har mer än 5 års erfarenhet av området och kan erbjuda inte bara kontroll av förändringar i dokument, utan även möjlighet att kontrollera dessa förändringar tillbaka mot originalet. DeltaXML Core 3.0 som ger snabbare integration och spridning, är företagets nyhet på Online. Återigen ett företag som har många av sina kunder bland övriga utställare.

/Gunnel



### ***EBSCO Information Services***

EBSCO har en aktiv roll i utvecklingen av en sk Web service för enkel hantering av statistik om tidskriftsanvändningen vid t ex bibliotek. Den utgör en del inom SUSHI (Standardized Usage Statistics Harvesting Initiative) och är framtagna i samarbete med NISO (the National Information Standards Organization) och COUNTER-projektet (som tar fram standarder för mätning av onlineanvändningen av informationsprodukter och –tjänster). SUSHI lanseras för test i början av 2006.

Forbes rankade i november 2005 EBSCO till plats 162 (av 339) på årets lista över största privatägda amerikanska företag.

/Gunnel

### ***Exegenix***

Exegenix är framstående vad gäller att konvertera information till XML, vilket man gör ”snabbt, korrekt, och kostnadseffektivt, oberoende av dess struktur”. Man säger också att deras teknik inte kräver någon mappning eller programmering men ändå genererar konsekvent strukturerad information. Den som vill se detta med egna ögon kan få testa gratis konvertering till XML via [www.exegenix.com/solutions/sample/index.html?enter=cme](http://www.exegenix.com/solutions/sample/index.html?enter=cme). Exegenix har bl a svarat för tekniken beträffande LexisNexis Butterworths publicering av den brittiska regeringens budget för 2005.

/Gunnel

### ***IEE /Inspec***

Alla löpande IEE publikationer finns nu tillgängliga online i IEE Digital Library. IEE tidskrifter från 1872-1993 håller också på att läggas upp. Arkivet 1898-1968 finns nu tillgängligt från de vanliga källorna. IEE kommer också med två nya titlar: Information Security and Intelligent Transport Systems.

/Gunnel

### ***IEEE***

I den samordnade tjänsten IEEE/IEE Electronic Library finns material från IEEE journals, transactions, magazines and letters, IEEE conferences, IEEE standards och IEE journals & conferences. Under 2006 kommer man att lansera en tjänst där användarstatistik kommer att visas. Man kan då följa hur en källas används, t ex per titel och antal sökningar.

Under 2005 har IEEE lagt in Proceedings of the IEEE från och med 1963, och under 2006 kommer man att följa efter med material ända tillbaka 1913. IEEE är och förblir en viktig källa för våra tekniskt inriktade bibliotekskunder, och källans ålder har vid dessa sökningar aldrig skrämmt mig.

Bland korta fakta om IEEE saxar jag även följande: IEEE är världens största tekniska yrkesförening med mer än 365 000 medlemmar i ca 150 länder. Inom flera av sina discipliner står IEEE för 30% av världens totala utgivning. IEEE har tagit fram mer än 900 aktiva industristandarder. Man är också med och sponsrar mer än 350 internationella tekniska konferenser varje år.

/Gunnel

### ***Infinity-loop***

Infinity-loop, som är ett tyskt XML-företag, visade under Online 2005 en nyhet exklusivt för dess besökare. Inom mjukvaruområdet för att konvertera dokument till XML har man en Java-baserad produkt; upCast©. Här kan man omvandla såväl Word- som RTF-dokument till XML, och tillbaka igen, fast då heter det downCast©. Infinity-loop har många kunder just bland utställarna på Online, bland annat LexisNexis och Elsevier.

/Gunnel

### ***Ingenta***

Ingenta presentade en ny Gateway Service för bibliotek som kallas IngentaConnect Complete. Den erbjuder biblioteken möjlighet att erbjuda sina användare åtkomst via ett gränssnitt som de själva utformar (med egna design etc) och current awareness alerts mm. Man kan komma åt fulltext via prenumeration eller pay-per-view. Prissättningen börjar vid 40.000 kr

/Gunnel

### ***Oxford University Press***

I Oxfords monter hade man en mycket slö dator och tyvärr blev inte mycket visat. Likafullt är man i full färd med att ta världen med storm inom områden som ePublishing Solutions and Services. Visst följer den mänskliga faktorn med ibland även på mässor och utställningar, men uttryck som det man använt i presentationsmaterialet "Could any publisher be more in harmony with the goals of librarianship?" landar kanske inte riktigt där man avsett.

Oxford lanserar The Oxford Digital Reference Shelf som sägs erbjuda biblioteken att köpa elektroniska utgåvor av många av deras publikationer. Vidare introducerar man en ny inköpsmodul för Oxford Scholarship Online.

/Gunnel

### ***Scientific, Technical & Medical***

#### ***CSA***

CSA hade ingen egentlig ny lansering. Men självklart framhöll man den tidigare under 2005 lanserade CSA MultiSearch för samsökning i multipla källor via ett gränssnitt. Den kan anslutas till 2.000 olika elektroniska "resurser". Och så berättade man om en kommande databas, nämligen den franska databasen FRANCIS från INIST-CNRS som inom kort kommer att bli tillgänglig i CSA Illumina. FRANCIS är en bibliografisk databas inriktad på humaniora och samhällsvetenskap. Den täcker över 4.000 tidskrifter men också en del böcker mm. Totalt ingår 1.6 miljoner referenser med abstracts, flertalet på engelska. CSA framhöll speciellt FRANCIS goda täckning vad gäller europeiska publikationer.

Jag räknade det totala antalet databaser i CSA Illumina till ca 110. STN erbjuder ett flertal av CSAs egna databaser och under det senaste året har ett antal av CSAs egna databaser (åter) blivit tillgängliga vid Thomson Dialog. Bland CSAs kommande produkter nämndes även nya områden inom "serien" Scholar Universe. Scholar Universe har nyligen startats med Scholar Universe: Social Science, som tillhandahåller uppgifter om över 100.000 forskare mm och 1.500 organisationer inom området, dock endast i Nordamerika. Man arbetar nu med Scholar Universe: Health, Scholar Universe: Technology och Scholar Universe: Life Sciences.



Strax efter Online meddelade CSA att man övertar ansvaret för publicering och distribution av alla Ulrich's produkter från sitt systerföretag R.R. Bowker.

CSA uppgavs ha ca 300 anställda i USA och 30 i UK.

/Lars

#### ***Elsevier Science: Scopus, ScienceDirect m fl***

Elseviers stora montergrupp upplevdes nästan som en ren kopia av förra årets. Samma placering, närapå samma utseende och samma huvudnummer, nämligen Scopus och ScienceDirect. Men varför ändra på ett vinnande lag? Scopus, som då var nästan ny, har nu etablerat sig som en av de viktigaste resurserna för bibliografisk informationssökning inom vetenskap och teknik.

Att slå ihop ett antal databaser till en enda, stor, databas men där användarna tack vare automatisk kategorisering av träffarna på olika discipliner (mm) slipper uppleva "google"-känslan av att inte ha kontroll över vad det egentligen är man söker i och vad man får träff på var uppenbarligen rätt recept

för framgång. Ty framgång tycks det ju ha blivit även om man inte kan vara alldeles säker eftersom Elsevier Science inte har särredovisat Scopus ekonomiskt.

Tidigare under 2005 uppgav Elsevier officiellt att Scopus har över 500 kunder, de flesta storkunder inom universitets- och den offentliga sektorn. Hur många är de nu då? Ja, på Online uppgavs över 1.000. Det är intressant att jämföra detta med att antalet organisationer som använder ScienceDirect är över 3.000. Detta enligt en intervju i IWR Sept 2005 med Amanda Spiteri, marknadschef för Elseviers bibliografiska databaser. Och i montern uppfattade jag – med reservation för missuppfattning - uppgifter om att därmed sådär 10 miljoner forskare, studenter etc har potentiell tillgång till tjänsten, att 5.000 användare är online ”samtidigt” och att tillväxten sedan 2004 har varit 48 %. ScienceDirect tillhandahåller numera ca 7 miljoner artiklar i fulltext från 2.000 tidskrifter.

En viktig faktor i framgången för Scopus är helt klart satsningarna på att involvera representanter för avnämarna i dess utveckling – inför lanseringen såväl som i samband med vidareutvecklingen.



*Utställningshallen med Elsevier Science och Scopus i förgrunden*

Personligen tror jag att det viktigaste har varit ”förankringseffekten”/marknadsföringseffekten snarare än produktens utveckling i sig. Inför lanseringen engagerades 300 användare och bibliotekarier i utvecklingen. Och nu vid Online 2005 gjorde Elsevier Science ett stort nummer av dess nya Content Selection Committee för Scopus, vilken består av 20 av ”världens främsta” vetenskapsmän och 10 ämneskunniga bibliotekarier, vilka ska tillförsäkra Scopus en ”unbiased and comprehensive research coverage”. Kommittén sammanträdde för första gången i Budapest i oktober 2005 och uppges vara den första ”oberoende kommitté” någonsin som engagerats i utvecklingen av en bibliografisk databas. I en artikel i Econtent October 2005 uppges att kommittén bl a har till uppgift att bredda innehållet i Scopus mer ”åt Google-hållet” men utan att ge avkall på den goda kvalitét som manuellt utvalt innehåll innebär. Scopus erhöll för övrigt utmärkelsen ”Best STM Information Product” vid Onlines ”Awards Ceremony”.

I övrigt berättades på Online 2005 om det senaste halvårets Scopus-nyheter, som t ex avtalet i september med JISC (Joint Information Systems Committee) som redan efter ett par veckor inneburit att 70 brittiska universitet börjat nyttja Scopus, att Scopus nu länkar till fulltext i artiklar i över 250 tidskrifter i JSTOR, att integration gjorts med CSAs RefWorks för samordnad och enkel redigering av referenser och att integrerad åtkomst till CrossFire Beilstein för kemiska strukturer erbjuds för dess kunder. Scopus refererar till totalt 14.000 tidskrifter från 4.000 förlag. Det synes fö som om man i stor utsträckning nyttjar artikelförfattarnas egna abstracts snarare än det traditionella förfarandet vid produktionen av bibliografiska databaser, dvs att låta egna redaktörer författa dessa.

Apropå samarbete med onlinetjänster utanför Elsevier rapporterades att den relativt nya databasen EMBiology nu gjorts tillgänglig via Ovid. EMBiology innehåller över 4 miljoner referenser från 1980. Därtill lanserades en betaversion av den nya databasen EMCare, som innehåller 2 miljoner referenser inom sjukvård, hälsa och biomedicin från 2.700 tidskrifter.

Förutom de nämnda produkterna förevisade Elsevier Science också Embase.com, Engineering Village 2, MDL's Discovery Gate, Endeavor's LinkFinder Plus och Scirus.

/Lars



## EU-kommissionen

EU-kommissionen hade som vanligt en av de allra största monterarna – i år även vad gäller ”höjd”. Där förevisades en mängd produkter och tjänster – som vi dock hann inte med att granska. Jag hann i alla fall ta ett foto ...

/Lars

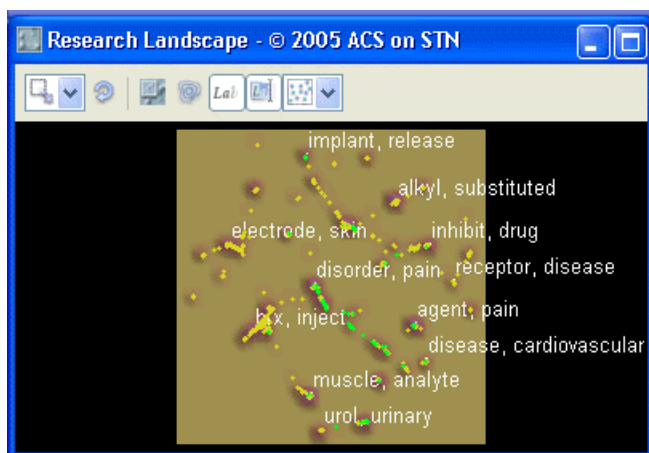
## FIZ Karlsruhe/STN International, FIZ Chemie, CAS (Chemical Abstracts Service)

STN AnaVist var huvudnumret i såväl CAS monter som i den monter där FIZ Karlsruhe/STN International och FIZ Chemie huserade. Typiskt nog för Online 2005 var det alltså ingen nyhet man framhävde: STN AnaVist lanserades ju redan i våras och har sedan dess marknadsförts i många sammanhang. Produkten utgör också en exponent för den trend som yttrar sig i att



En del av EU-kommissionens monter

onlinetjänsterna mer och mer satsar på verktyg mm för analyser av information och visualisering medan ny information kommer i bakgrunden.



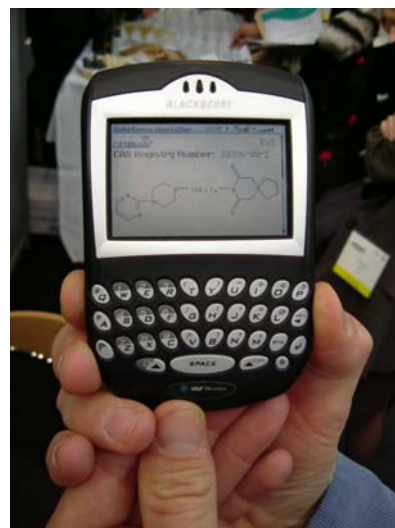
STN AnaVist ger möjlighet att identifiera och åskådliggöra trender genom att analysera sökresultat och uppgifter från litteratur och patent, funna i databaserna CAPlus, USPATFULL och PCTFULL. Sökresultat på upp till 20.000 dokument kan analyseras. Ett exempel på visualisering visas i figuren, där varje punkt representerar ett funnet dokument. Ju tätare punkterna förekommer desto mer forskningsintensivt är området.

FIZ Karlsruhe/STN International berättade också om att den nya versionen av Derwent World Patents Index (DWPI) kommer att bli tillgänglig via STN under första kvartalet 2006. Databasen är anpassad till IPCs nya klassificering IPC8 som träder i kraft den 1 januari. Därtill är ytterligare 750.000 abstracts sökbara i fulltext och samtidig vänster- och högertrunkering medges för sökning i flertalet fält. Man presenterade också den nya versionen av FIZ AutoDoc, där en av nyheterna är möjlighet att beställa dokument pay-per-view från ett antal förlag. Stöd finns för Open URL 1.0.

FIZ Chemie presenterade databasen Infotherm, gratis sökbar via [www.chemistry.de/infotherm](http://www.chemistry.de/infotherm). Infotherm tillhandahåller fakta om 6.300 ämnen och 23.000 blandningar. Data för flertalet ämnen ges utan kostnad medan blandningar kan köpas via pay-per-view eller via abonnemang.

CAS lanserade CAS Mobile, vilket som namnet antyder rör mobil åtkomst till CAS databas. Det som förevisades var åtkomst via Blackberry. CAS Mobile kan nyttjas utan kostnad av storkunder till SciFinder och STN-kunder med fastprisavtal.

/Lars



(Foto: [www.infotodayblog.com](http://www.infotodayblog.com))

### ***MyiLibrary***

MyiLibrary var en av de nya aktörerna på Online 2005. MyiLibrary, som lanserades i slutet av 2004, är en webbaserad ”plattform” för sökning i olika e-bokssamlingar, bl a från Wiley, Springer, Elsevier, OUP, ILO, WHO, Världsbanken, OECD m fl. Svenska Omnistat ingår också bland leverantörerna. Plattformen, som i sig kostar 2-5.000 kr per år, tillhandahålls via ProQuest och systerföretaget Coutts Library Services och den riktar sig till bibliotek inom den akademiska världen och offentliga sektorn. Jag fick uppfattningen att man ännu inte har någon svensk kund.

/Lars

### ***ProQuest***

ProQuest tillhandahåller en bred flora av produkter och tjänster (ProQuest, Chadwyck-Healey, UMI, eLibary, SIRS/Voyager, XanEdu, ...). Den person jag pratade med i montern berättade att man tillsammans med sitt systerbolag Serials Solutions arbetar på att göra så mycket som möjligt av allt som ingår sökbart via en och samma ingång och gränssnitt (”federated search”), med användning av sin produkt Central Search.

På Online 2005 presenterade ProQuest bl a sitt samarbete med MyiLibrary, som går ut på att via MyiLibrarys plattform tillhandahålla drygt 2.000 e-böcker inom medicin mm från ProQuest, och sitt omtalade Nineteenth Century Index (C19) med 11 miljoner bibliografiska poster till böcker, tidskrifts- och tidningsartiklar mm.

ProQuest nylanserade också sina databaser PCI och PCI Fulltext, vilka efter två års arbete nu tillhandahålls via ett helt nytt gränssnitt och under nya namn, nämligen Periodicals Index Online respektive Periodicals Archive Online. Därtill framhöll ProQuest sin produktsvit inom affärs- och företagsinformation, där förutom den egna databasen ABI/Inform ingår bl a Factiva (som ju inte självt ställde ut på Online), marknadsrapporter från Snapdata, företagsprofiler från Hoover’s, e-böcker från Safari, Wall Street Journal och Oxford Analytica.

/Lars

### ***Questel-Orbit***

I franska Questel-Orbits monter möttes jag av öppna famnen av veteranen Pierre Buffet som jag ju träffat med jämna mellanrum ända sedan min tid vid IDC-KTHB i slutet av 1970-talet (– att jag möttes av hälsningsfrasen ”Trevligt att träffa en annan dinosaur!” fick jag stå ut med). Pierre är en av grundarna av Questel-Orbit (och vid Online 2001 erhöll han Online Information Lifetime Achievement Award). Pierre berättade att Questel-Orbit fortsätter att satsa i sin nisch, dvs patent och varumärken, och att man inte har några planer på att ge sig in på andra områden.

Det går bra för Questel-Orbit (som också säljer bra i USA, där man har en filial), men givetvis utsätts man från hård press från gratisinformation på webben och den största konkurrenten Thomson (fast Questel-Orbit tillhandahåller ju fortfarande dess Derwent World Patents Index ...). Pierre berättade att även LexisNexis med sina satsningar alltmer börjar framstå som en konkurrent. Precis i enlighet med branschen i övrigt är nysatsningarna inriktade på analysverktyg av olika slag snarare än ny information. Man förevisade på Online 2005 produkter som finns sedan tidigare, bl a ett verktyg för att ”hitta på” nya varumärken och PatentExaminer för analyser av funna varumärken, såväl som helt nya CaTaMaran Plus.

CaTaMaran Plus har utvecklats av Edital, ett företag som förvärvats av Questel-Orbit. Produkten erbjuder djupanalyser av varumärken och varumärkesinnehavare, bl a likhetsanalys, redovisning av nya varumärkesansökningar och varumärkestsvister, konkurrenternas varumärkesinnehav mm. CaTaMaran baseras på varumärken utfärdade enligt Madrid-överenskommelsen och av EU, UK, Tyskland, Frankrike och Benelux.

/Lars



### **Thomson Scientific**

Thomson Scientifics monter förde i första hand fram Derwent, som man berättade kommer att lanseras i en ny version under första kvartalet 2006. Nytt är då bl a en anpassning till nya patentklassificeringssystemet IPC v. 8 och möjlighet till sökning även i texten till 750.000 patentsammanfattningar. Den nya versionen blir tillgänglig på såväl Dialog som STN och Questel-Orbit. Därtill berättade man om den kommande Thomson Data Analyzer, som är en ny version av Derwent Analyzer. (Ja, det är populärt med analysverktyg ...).

I montern marknadsfördes också ISI Web of Knowledge, Thomson Pharma, MDC och många fler produkter och tjänster. "Gamla" Dialog då? Jodå, den fanns också representerad men var ganska så nedtonad jämfört med vad man är van vid - och jag hade faktiskt lite svårt att få information om tjänsten i montern (för att inte tala om minen när jag frågade om TBI, dvs den andra delen av Dialog, vars monter jag letat efter i utställningsguiden men inte funnit eftersom den var listad under Business Intelligence Services!).

/Lars



*Thomson Scientific och Dialog (t.v.) i skuggan av Thomson Gale*

### **Wolters Kluwer: Ovid, Silverplatter m fl**

Wolters Kluwer upprepade förra årets stora satsning på Online med en av utställningens största montrar. Men det man som besökare noterade mest var nog dess nya logotyp. Den symboliserar att Wolters Kluwers visserligen består av ett antal delar men att de hålls samman (cirkeln). Som en följd är Ovids egen logotyp nu borta – men Ovid är ändå fortsatt profilerat som en egen produkt; faktiskt den enda av de jag såg i monterna som hade sitt namn snarare än Wolters Kluwers i anslutning till logotypen.



Ovid uppmärksammade förstås sina senaste lanseringar innan Online. De var Ovid Lingua (Ovids "multi-language" plattform som erbjuder sökning på 80 språk), QUOSA Information Manager (en Ovid-anknuten produkt för lokal, PC-baserad, hantering - analyser, redigering mm - av sökresultat som tagits fram av programvaruföretaget QUOSA) samt den utvidgning till 350 e-böcker i Books@Ovid man gjort i och med ett par nya förlagsavtal.

Apropå Ovid så syntes varumärket Silverplatter inte alls, förutom att det var nämnt i utställningsguiden, varför jag precis som jag "alltid" gör ställde frågan om dess framtid. Svaret var – också precis som alltid – att produkten är vid full vigör och ska bestå. Och det finns fortfarande många kunder till CD-ROM-produkterna.

Wolters Kluwer Health lanserade en ny produkt, benämnd PubFusion. Den riktar sig till den akademiska förlagsindustrin och utgör en komplett plattform för publicering av tidskrifter och utgör en del i EMCs produkt EMC Documentum 5 Digital Asset Management. Den togs ursprungligen fram för användning för publicering av 160 medicinska tidskrifter inom Lippincott Williams & Wilkins och används nu för totalt 250 tidskrifter internt inom Wolters Kluwer. PubFusion kommer att finnas till försäljning under 2006.

Hur är det då med Wolters produkter och tjänster inom området rättsinformation? Ja, inte märktes de i alla fall mycket på årets Online. Och ingen av dem jag träffade i montern kunde heller ge någon avsevärd information om dessa. Fast av de svar jag fick tolkade jag att den storvulna satsning på en gemensam plattform för hela koncernen, internationellt, för såväl publicering som sökning av rättsinforma-

tion, benämnd ”Single Internet Platform”, kommit av sig. En delorsak skulle i så fall kunna vara att de påtänkta NXT-programvarorna bytte ägare förra hösten – från Nextpage till FAST.

Vid sidan av Online pågick - och pågår - envisa rykten om att Wolters Kluwer för samtal med potentiella köpare, Thomson och Reed Elsevier, något som dock ihärdigt förnekas av ledningen. Bakom ryktena finns bl a tveksamheter om Wolters Kluwers förmåga att klara sig ekonomiskt på egen hand.  
/Lars

## 5. Föredragen

Konferensen öppnades som sig bör av konferensens ordförande **Martin White**, Intranet Focus. Innan han lämnade ordet till konferensens keynote-talare gjorde han en liten tillbakablick: 1965 - Dialog var under utveckling, 1975 - realtidsuppdatering, 1985 - disketter, 1995 - den 12 december, dvs för nästan exakt tio år sedan, hade Altavista 16 miljoner webbsidor. Nu, år 2005 - har Google 8 (eller är det 12?) miljarder webbsidor. Hur är det med oss informationsspecialister? Har vi utvecklats i samma mån? Det ville Martin låta vara osagt - men han lovade att Online 2005 i alla fall skulle bidra.



### *The new shape of knowledge – everything is miscellaneous*

Konferensledningen engagerar alltid en namnkunnig person för konferensens inledande keynote-föredrag. I år var det **David Weinberg**, känd som medförfattare till boken "The Cluetrain Manifesto – the end of business as usual". Han är ledamot av styrelsen för Berkman Center for Internet and Society vid Harvard Law School och jag har hört honom kallas för "en av världens mest inflytelserika tänkare om framtiden för Internet". Titeln på hans keynote på Online var "The new shape of knowledge – everything is miscellaneous".

David's föredrag utgjorde ett angrepp på "vårt" hittillsvarande sätt att organisera kunskap (sortering, klassificering, metadata etc, utförd av producenten själv eller av experter, t ex bibliotekarier), något han menade är fåfängt och meningslöst i det framväxande nätverkssamhället, där all kunskap sammanbinds av de som agerar där; dvs de som "använder" kunskapen (t ex länkar till den). "Knowledge is conversation. Knowledge is the work of human connectors". Sociala nätverk är framtiden!

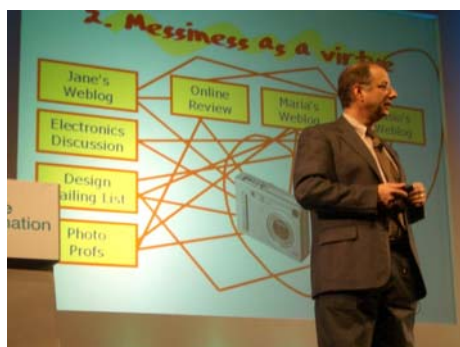
Tillsammans förfogar vi människor över en gemensam kunskapsmassa. Problemet är bara att kunskap inte är något konkret och väldefinierat utan mångtydigt och oklart. Och ingen kunskap är fullständig.



Därtill är den komplex. Hur har människan då av tradition tagit sig an kunskap? Jo, hon har försökt organisera den, "komprimera" den - för att på så sätt få översikt, kunna skapa sig kontroll. Hierarkiska modeller, klassificeringssystem, kortkataloger med möjlighet till multipel kategorisering, metadata ... Det är en mänsklig mani! ("Bibliotekarierna blir inte bibliotekarier pga att det ger pengar utan pga att de gillar att organisera kunskap").

Men problemet är att detta inte fungerar i längden. Hur man än förfinar indelningen, utökar metadatat etc så finns det aldrig något slut; se bara på hur Dewey-klassificeringen utvecklats. Det går aldrig att fullständigt beskriva något –

inte minst därför att vi ju inte vet hur det kan komma att nyttjas i framtiden!



Se t ex på hur ”kunskapen” om en kamera - eller ett foto - kan användas i dagens nätverksamhälle. Den/det kan nyttjas i hur många sammanhang som helst! (Se figur!) Det blir en stor ”messiness”; en ”unknown order”! Väsentligt här är att det förut var ägaren som ägde sorteringen – men det är det inte längre! *Det är användarna som svarar för – och därmed äger - ”sorteringen”/ordningen!*

Visst kommer det traditionella ”trädet”, den hierarkiska trädstrukturen, att fortsätta att vara användbar. Men inte ett – utan i mängder, i ett stort antal! Föreställningen om att upphovsmannen/ägaren/bibliotekarien ska vara ett ”filter” för andra är obsolet! Det var relevant i papperssamhället; ett papper kräver val av ”fack”; men nätbaserad kunskap kan ”sorteras” i hur många ”fack” som helst – och av vem som helst för vilket ändamål som helst. Idag eller om hundra år.

Det har länge sagts att vi kommer att drunkna i information. Men det har inte hänt, och kommer aldrig att hända – ty det finns ingen gräns vad gäller metadata och vad vi anbringar den på: trädet, grenarna eller löven. Eller om det beskriver ”betydelse” eller ”personer” eller vad som helst. ”It’s an analog world”, dvs ”messy” – trots att tekniken är digital!

David menade att Wikipedia är ett av de mest betydelsefulla uttrycket för denna nya värld – just för att den inte är pappersbunden. ”Jämför med Encyclopedia Britannica, som består av 22 volymer. Varför just 22? Det är för dyrt med fler! Alltså: ’Economics of paper drives knowledge’. Knappast vettigt eller hur?” Wikipedia kan däremot utvidgas i oändlighet.

Bloggar är en parallell; de bygger ju på att man relaterar till andra; länkar till andra. ”Utan länkar – ingen webb”. *Bloggare är generösa!* Detta ställde David i kontrast till kommersiella media, typ New York Times, vars länkar bara går ”utåt”. ”NYT är narcissister”. De är rädda för att folk ska lämna dem, dvs inte betala! Trovärdigheten då? Tja, ”varför ska man tro mer på en redaktör i NYT än på Jimmy Wales” (Wikipedias grundare)? Objektivitet skapar vi själva. David använde ordet ”*multisubjectivity*” och tog som exempel hur olika ett och samma foto kan uppfattas av olika personer.

Efter dessa synpunkter var det inte alls svårt att gissa sig till att David hade en positiv syn på Google och andra webbens sökmotorer, trots dessas brister. ”Anta att jag beställer in en öl, en lager, på en bar jag aldrig besökt förut. Jag får min lager och konstaterar att den inte är perfekt. Men den är tillräckligt bra. Samma sak är det med Google-sökningar: de är inte perfekta men de är *tillräckligt* bra!” Och i själva verket kan man ju faktiskt uppfatta Googles träfflistor som en funktion av ”human connectors”.

## ***Paneldiskussion - Desktop search***

Därpå delades konferensen upp i tre parallella spår. Jag valde den som var inriktad mot ”sökning”, vilken jag följde resten av dagen (de övriga två var inriktade mot bibliotek och publicering respektive CM). Den första sessionen utgjordes av en paneldiskussion med titeln ”Desktop search tools in the spotlight”. Den leddes av sökmotorgurun och tillika veteranen i Online-sammanhang **Stephen Arnold**, Arnold Information Technologies (USA). Deltagare var i övrigt **David Graham**, produktansvarig vid MSN (UK), **Franck Poisson**, Exalead (Frankrike), **Suranga Chandratillake**, Blinkx grundare (UK; har tidigare arbetat vid Autonomy) och **Nikhil Bhatla**, produktansvarig Desktop Search vid Google (USA).

Orsaken till att jag valde denna session var inte att jag är speciellt intresserad av just ”desktop search” utan för att jag är intresserad av sökmotorer (”enterprise search engines”) i allmänhet. Och Stephens första fråga till panelen rörde just detta, nämligen ”vad är det för skillnad mellan desktop- och enter-

prise-sökmotorer”? David svarade att de är mycket lika och relaterade till varandra; vilket också är orsaken till att Microsoft nyligen också lanserade en enterprise-version. Franck menade att det inte är någon skillnad ur användarnas synpunkt; gränssnittet är detsamma och det spelar ingen roll vad som finns ”under”. Suranga och Nikhil betonade dock att lokala sökmotorer har fördelar i det att de kan åstadkomma verkligt effektiv lokal sökning samtidigt som de utgör gränssnitt mot ”stora världen”.

Vilka skillnader finns då mellan de aktuella produkterna? Självklart omöjlig att svara kort på – men följande sades. Microsofts produkter utgör integrerade delar av Windows. Karakteristiskt är att samma teknologi finns bakom i såväl desktop- som enterpriseversionerna och att applikationerna är mycket ”öppna”. Exalead hanterar alla format utan att användarna behöver bry sig. Blinks håller reda på vad den enskilde användaren gör och ger anpassad hjälp. Google slutligen sparar kopior av allt som visas på skärmen (e-mail, webbsidor, worddokument ...), vilket inte bara gör det möjligt att enkelt återfinna sådant man minns att man arbetade med/tittade på ”härnäst” utan även återskapa sådant man kanske missat att spara eller tappat bort.



Fr v: Stephen Arnold, David Graham, Franck Poisson, Suranga Chandratillake, Nikhil Bhatla

Apropå integration: Vad händer på marknaden om Microsoft och Apple gör sökning till en integrerad del i sina operativsystem (vilket en i publiken följande menade inte var en fråga om ”om” utan om ”när”)? ”Det spelar ingen roll”, svarade Suranga kaxigt, ”ty de kommer att klara filhantering m fl systemfunktioner men däremot inte bli fullödigas söksystem”. Franck trodde för den delen inte alls på att de kan få monopol på desktop, inte minst därför att det alltid kommer att finnas en skillnad mellan användning av produkter för allmänheten och för företagen.

Det kommer alltid att finnas konkurrens! (Anm: jag noterade här att panelens deltagare uppenbarligen betraktade desktop som något som enbart används privat mellan ”enterprise search” rör företag. Det sista är såklart givet men knappast det första ...). David framhöll att sökning visserligen kommer att integreras men att Microsoft för sin del är mycket noga med att inte låsa in lösningarna; de kommer alltid att hållas öppna mot omvärlden genom att APIer alltid kommer att tillhandahållas som underlättar kopplingar till omvärlden.

En av de frågor som luftades i det följande diskussionen var frågan om hantering av dokument i multipla versioner. Går det att skilja på dem eller indexeras allt? Här upplevde jag svaren lite svävande. Blinks tycks i alla fall kunna upptäcka rena duplikat. Google visar alla, men ”de går att ta bort”. I Exalead kan administratören styra vem som ska se vad. Även frågor kring hur produkterna fungerar i stora nätverk, typ serverfarmer, gavs svävande svar, liksom frågor kring säkerhet/åtkomst. Uppenbarligen saker att vara observant på inför ev. anskaffning! Allra mest luddiga svar gavs följande betr hur uppdatering i index av nytilkomna dokumenten sker (t ex hur snabbt). Självklart är detta beroende på belastningen i övrigt (t ex samtidiga sökningar), serverstorlek etc – men det tycks vara en akilleshäla.

Själv ställde jag en fråga om hur panelen trodde att marknaden för desktop-sökning skulle utvecklas. Detta med anledning av att det ju idag finns en uppsjö av produkter att välja på. Svaret blev att jo, den kommer att konsolideras – men att det inte kommer att bli frågan om några monopol. Och nischprodukter finns det alltid plats för. Beträffande produkter för konsumentmarknaden kommer det alltid skapas nya sådana: att bygga söksystem är ju så enkelt.

En annan fråga rörde Googles principer för relevansbedömning. Nikhil svarade att dess desktopversion skiljer sig avsevärt från ”stora” Google i detta avseende. T ex nyttjas inte ”page rank”. Stor hänsyn tas däremot till när dokumentet sist användes. Microsoft betonade möjligheten att styra relevansen m hj av olika viktning av fält och Blinks berättade om ”filter” för olika typer av dokument.



## *Search is no longer about search*

Den andra sessionen i ”sökspåret” var sessionen ”Tools for effective information management”, med webbsökmotorgurur **Chris Sherman** som moderator.

Den inleddes med ett föredrag av **Phil Bradley**, Internet Consultant, UK. Det hade titeln ”Search is no longer about search”. Han inledde med att påpeka att det numera inte är vi som behöver sökmotorer – det är de som behöver oss. Och de jagar oss med alla medel: Google satsar på toolbars, nyhetsgrupper, Gicasa, Gmail, desktop-sökning, boksökning, sökning i vetenskapliga tidskrifter etc etc – allt för att dra oss till sig (MSN gör på motsvarande sätt). Google är ingen webbsökmotor – det är ett företag! Sökning blir alltmer ”mobil”; och vi söker folk, multimedia, platser. Snart finns det chips i ”allt” – och allt kommer att förenas. Lokalisera sina vänner, söka dokument i sin PC, hålla sig uppdaterad, aldrig gå vilse, lära känna sina omgivningar, ... allt. A9 har 35 miljoner foton över USA. Och var och en av oss utgör ”informationsresurser”! Allt kan digitaliseras – och sökning blir alltmer ”osynlig” i sig; den blir en självklar komponent i vårt dagliga liv. Sökning är inte sökning!

## *Comparison of desktop tools*

**Karen Blakeman**, RBA Information Services, UK, välkänd för sitt nyhetsbrev Tales from the Terminal Room (TffTR), presenterade ”Comparison of desktop tools”. Hennes presentation finns på [www.rba.co.uk/presentations](http://www.rba.co.uk/presentations) men här några ”highlights”.

Desktop-sökning innebär att indexera och söka material som finns i den egna PC:n: dokument i Word, Excel, Powerpoint, chatt, IM, e-mail, webbsidor etc etc. En del produkter kombinerar lokal sökning med sökning på webben. Viktiga saker att beakta vid val av produkt är vilka dokumentformat som stöds, hur det går att styra indexering till mappar och format, minnesanvändning vid indexering liksom tid för indexering (går det att pausa?), sökfunktioner, möjligheter till förhandsgranskning, användarvänlighet, säkerhet och auktorisation.

Precis som i paneldiskussionen om desktop-sökning blev indexering speciellt omnämnd – det är uppenbarligen en kritisk punkt. Viktiga saker att beakta framgår av den bild som Karen visade. Se fotot här intill!

Resten av föredraget ägnades åt en genomgång av sju produkter, nämligen Ask Jeeves, Blinkx, Copernic, Exalead OneDesktop, Google, MSN och Yahoo.

**Ask Jeeves** var initialt en ”katastrof”, men är nu bättre. Speciellt videosökning är bra. Inte en ”ideal” produkt – men den är helt OK.

**Blinkx** var i tidigare versioner svåränvänd men har nu blivit bättre. Klarar över 200 filformat – dock inte Openoffice. Innehåller en visualiseringsfunktion som verkar intressant!

**Copernic** har funnits mycket längre än de andra produkterna. Den finner även udda filformat. Medger boolesk sökning och avancerad parentessökning.



*Karen Blakeman i talarstolen. T.v. vid bordet Chris Sherman och bredvid honom nästa talare, Mark Sanderson*



**Exalead** har allt som webbsökversionen har ("så älskar du den så kommer du att älska denna"). Den klarar "approximativ" sökning, maskning, närhetssökning etc. Största bristen är att den inte klarar alla format, t ex Openoffice – men det sägs vara på gång. Betyg: "Mycket trevlig produkt".

**Google** fick kritik av Karen. Det viktigaste skälet till detta är att kopior av borttagna dokument sparas. De kan visserligen tas bort separat - men funktionen är inte idiotsäker, vilket gör detta till en säkerhetsrisk. Karen hade egen erfarenhet av detta: Efter två månaders användning av Google upptäckte hon plötsligt att allt hon gjort under denna tid hade lagrats i Google utan att hon varit medveten om det. (Anm: Det här har jag hört ryktas om tidigare. Att det inte framkom i paneldiskussionen ovan - även om jag kände att det låg på lut - berodde kanske pga artighet mot Google-representanten där?). Ett annat skäl var att Google är resurskrävande, något som märks även vid användning av kraftfulla maskiner. Resursanvändningen vid indexering går inte heller att styra, vilket kan påverka samtidig sökning.

**MSN** sade Karen inte så många ord om. Jag fick i alla fall uppfattningen att hon betraktar produkten som ännu ganska primitiv och ofullgånge, något som kanske kan bero på att man satsar allt krut på kommande versioner av Windows och dess integrerade sökfunktion.

**Yahoo** sades inte heller så mycket om. Installationen kan tydligen vara problematisk. Bra är att den klarar många filformat, faktisk flest av alla.

Vilken desktop-sökprodukt ska man då välja? Det beror på vad man ska använda den till. Karen gav alltså inget svar – men jag uppfattade att hon gillade Exalead bäst. Troligen använder hon dock en annan själv eftersom Exalead inte klarar alla de format som hon använder.

I samband med avtackningen tog Chris tillfället i akt att inventera åhörarnas användning av desktop-sökverktyg. Faktiskt räckte så många som mellan hälften och två tredjedelar av publiken (ca 100 personer) upp handen.

## *Tools developed for image retrieval*

Det sista föredraget i sessionen bestod oss **Mark Sanderson** från Information Studies vid Sheffield University med. Det hette "Tools developed for image retrieval" och av det hade jag förväntat mig en fascinerande översikt över vad som går att göra (i stil med ett föredrag jag åhörde vid Online 2004). Men jag blev besviken. Kanske med rätta? Ty om man ska tro det Mark sade så finns det faktiskt inga bra söksystem för bilder idag! (Ja, det jag sett i andra föredrag och läst om är nog mest fantasifulla skapelser – kul att höra om och se, men som bara finns i sinnevärlden).

Bildsökning idag består mest av sökning i de texter som bilderna är försedda med. Och det beror på de oerhörda svårigheterna att identifiera bara så synbarligen enkla saker som träd, hus mm. Hur ska systemen veta exakt vad på ett foto som ska beskrivas? Träd, mark, löv, stam, grenar, färger, byggnader ... – det finns ju oändligt mycket att identifiera och beskriva. Och varje sak i sig är inte enkel att urskilja. Är det frågan om 3D-foton förbättras situationen: då går det ofta att identifiera och rikta in sig på näraliggande föremål.

Ansiktsgenkänning är ett område det bedrivs mycket arbete inom. Det fungerar hyfsat – men är ändå inte så pass bra att det kan nyttjas i praktiken för att t ex identifiera efterlysta personer. För att vara ens tillnärmelsevis pålitlig kräver automatisk igenkänning en exakt vinkel vid fotograferingen. Minsta avvikelse från "originalet" – och det fungerar inte. Undantag är personer som "kartlagts" och beskrivits i minsta detalj, typ kända personer som Tony Blair och George Bush.

Fast det finns trots allt avgränsade områden där bildsökning fungerar bra! Mark nämnde som exempel möjligheten att söka efter kända bilder i stora bildsamlingar (användbart t ex för att identifiera obehörig användning av upphovsrättsligt skyddat material), varumärkssökning (som fungerar hyfsat bra för abstrakta bilder och i alla fall spar en del manuellt arbete) och sökning i handskrivna texter.

## ***Paneldiskussion - Information discovery: your searching questions answered!***

Givetvis upprepades de senaste årens populära panelsession "Your searching questions answered!". Årets moderator var **Phil Bradley** och i panelen satt nästan bara namkunniga personer, nämligen **Chris Sherman**, Searchwise (USA), **Amelia Kassel**, MarketingBase (USA), **Mary Ellen Bates**, Bates Information Services (USA) och **Marydee Ojala**, redaktör för Online Magazine, (USA) (i nämnd ordning fr.v. på fotot).



På förmiddagen hade ju allt vad heter klassificering liksom andra sätt att försöka ordna upp eller beskriva kunskap fått ganska mycket stryk. Så Phils första fråga till panelen var om de trodde på en framtid för vokabulärer och uppmärkning, "tagging" (som var det uttryck man nyttjade så det gör jag också). Och här var meningarna delade! Marydee svarade med en motfråga: "Vem vill att allt ska hamna i 'miscellaneous'-boxen?" och svarade själv "Ingen!". Som exempel på behov berättade hon om den "spontana" kategorisering som uppstår

t ex i bloggvärlden. Så har t ex Online 2005 fått bloggtaggen "oi05" (Technorati)! Mary Ellen anförde att det där med att skapa taxonomier sitter djupt i oss och sade därför ja till tagging.

Chris däremot trodde inte alls på tagging. Visst, det fungerar jättebra. Men problemet är att 1) Folk är lata och de orkar inte tagga och 2) Det finns spammare! De kommer att förstöra på samma sätt som de förstör e-mail. Därför är det inte någon vits med tagging; sökmotorerna kommer helt enkelt inte att nyttja den. Christ slutsats var "Jag är synnerligen skeptisk till tagging". Han fortsatte med att säga att det är lönlöst att jobba med tagging när det ändå inte är det som avgör om sökmotorerna ska ta in det. Se t ex på Flickr, som bara indexerar ett fåtal av de bilder som faktiskt finns på webben. Och faktum är att "ingen sökmotor av idag nyttjar metadata". (Anm: Här drar jag mig till minnes tidigare års sessioner där jag ställt frågan om acceptansen för Dublin Core - och då fått svaret att den är mycket liten och att sökmotorerna inte alls beaktar den. Nu ansåg jag det inte ens lönt att fråga).

Marydee menade att det i alla fall är bra att tagga. Det kan ju komma till nytta i massa sammanhang, t ex i en speciellt utvald samling av information; kollektion. För sökning såväl som för att skapa alfabetiska listor eller för "tag clouding" (som jag förmodar i t ex visualiseringssyfte).

Nästa fråga från Phil var om det går att klara sig utan Google? "Ja, det gör det", svarade Mary Ellen omedelbart och fortsatte "för övrigt har folk som anlitar mig för sökuppdrag redan googlat". Även Marydee svarade ja "... och om jag verkligen behöver nyttja sökmotorer så använder jag Yahoo! först." Så berättade hon om att hon häromdagen träffade på en person som istället för att säga att hon skulle googla spontant nämnde Yahoo! vilket Marydee menade var "en nyttig erfarenhet!" och tillade att "till saken hör att hon inte är informationsspecialist".

"Leder Google till 'fördumning' vad gäller sökning?" löd så en fråga från publiken. Nej, inte alls, svarade Mary Ellen. Det är som med mat – alla äter inte nyttigt. Att det finns McDonald gör inte att alla äter hamburgare. Marydee tillade att sökmotorerna tvärtom har fått folk att söka mer avancerat. I en studie som CSA gjorde för tio år sedan visade det sig att de flesta bara använde ett sökord. Nu är det vanligare med flera. Chris instämde; antalet sökord i sökfrågorna ökar "... men det är ytterst få användare som nyttjar mer avancerade sökfunktioner".

Här kom diskussionen osökt in på sökordsoptimering, denna snabbt växande affärsverksamhet som går ut på att ”bra” sökord i brist på bra innehåll ska få sökmotorerna att indexera kundernas webbplatser. Må vara att denna verksamhet ofta går ut på att ”lura” sökmotorerna – den ger också sökmotorerna mycket erfarenhet om sökning. Apropos kommersiell verksamhet berättade Mary Ellen att hon numera ofta tar del även av sponsrade sökresultat eftersom de mycket väl kan innehålla relevanta träffar.

Videosökning – hur är det med det? undrade någon i publiken. Mary Ellen svarade att det inte används särskilt mycket. För övrigt – hur ofta är det man söker efter just rörliga bilder av något? ”Jag kan inte se att det någonsin kommer att bli stort vad gäller sökning”. För övrigt är det ju väldigt primitiva upptagningar som presenteras jämfört med ”riktiga” produktioner, menade Chris. En orsak är att video är väldigt krävande vad gäller datorkapacitet. Men Marydee bad oss inte underskatta nyttan med video. Text kan en videoupptagning av vad en finansanalytiker säger om text en storaffär avslöja betydligt mer än bara en textutskrift av hans tal – pga att det går att se på hans/hennes kroppspråk vad den egentliga innebörden kanske är.

Vad trodde då panelen om framtiden för kostnadsbelagd information och betaltjänster? Kommer den att försvinna? Nej, inte alls, blev det eniga svaret. Se bara på Yahoo! Subscriptions. Glöm inte att Yahoo! är ett ”innehållsföretag”. Jämför med situationen beträffande TV – folk är beredda att betala för att se det då vill ha. Se på framgången för iTunes. Sökmotormarknaden kommer att utvecklas på ett liknande sätt. Vi kommer att få se ”evolution of fee-based services”. Onlinetjänsterna har en lysande marknad att se fram emot på grund av ”information overload”, sade Chris. ”Folk har ont om tid. Jämför med hur marknaden för färdigmat har utvecklats! Framtiden är ljus för både gratis- och betaltjänster”.

Slutligen undrade Phil vad paneldeltagarna skulle göra om de fick basa för Internet under en dag. ”Införa närhetssökning”, svarade Marydee. Mary Ellen: ”Strypa spamarna”. Chris: ”Stänga Internet – tänk vad skönt att få vila från det en dag!”

## *China in the information economy*

Under andra dagen ägnades en hel session åt Kina; något unikt eftersom det var första gången någonsin i Onlines historia man ägnade så mycket uppmärksamhet åt ett enda land (tror jag i alla fall). Det inledande keynoteföredraget, ”STM Online Information Service Industry in China”, hölls av **Xiadong Qiao**, direktör vid Wanfang Data, Kina, tidigare 13 år vid Institute of Science and Technology (ISTIC). Wanfang Data, som var ISTICs IT-avdelning fram till 1993, är det ledande företaget i Kina inom STM, online och databaser och omsätter idag 125 miljoner kr.



Xiadongs dragning var synnerligen faktaspäckad men ändå ”redig”. Här några uppgifter från det föredraget (hans Powerpointbilder är tillgängliga via [www.wanfangdata.com.cn/ppt/on-line.ppt](http://www.wanfangdata.com.cn/ppt/on-line.ppt) - obs långsam server!) kombinerat med uppgifter från dokumentationen av det efterföljande föredraget ”Impact of globalization on Chinese information and publishing communities” som hölls av **Qiaoqiao Zhang**, chef vid China CAB International, UK, men som jag inte hann höra.

*Xiadong Qiao*

Kina satsar stora resurser på FoU inom vetenskap och teknik. 1998-2002 rörde det sig om ca 110 miljarder kr, vilket var mer än dubbelt så mycket som 1993-1997 och 2002 innebar 1.13 % av BNP. Under 2004 tycks en ombildning av statliga förlag till kommersiella ha skett i stor skala. Kina är en stor producent av böcker. 2003 publicerades 110.000 nya böcker, vilket kan jämföras med Storbritanniens 120.000 och USAs 175.000. Tidskrifter inom STM publiceras nästan uteslutande av universitet och vetenskapliga sällskap. Det finns omkring 10.000 tidskrifter och dagstidningar i Kina. 4.500 av dessa är inom STM, varav 210 på engelska. Totalt omsätter kinesiska tidskrifter ca 12 miljarder kr årligen, vilket gör Kina till en av de ledande i världen – men man ligger ändå efter flera

nationer i väst. Och dess tidskrifter inom STM har bara en liten roll inom vetenskaplig kommunikation, globalt sett.

Kina startade verksamhet inom online i mitten av 1980-talet. Det rörde då användning av internationella onlinetjänster, egna system för SDI och produktion av databaser. Internettjänster och fulltextdatabaser tillkom under 1990-talet. Så under 2000-talet har kommersiella online- och andra tjänster tillkommit. Sedan 1990 har ett antal informations- och dokumentationscentraler i Kina omvandlats till företag. För Wanfang Datas del skedde det 1993 och de publicerar nu ett antal databaser, bl a fulltextdatabaser som täcker 2.300 kinesiska tidskrifter inom STM. En annan stor producent av databaser är Chinese Academy of Sciences med sin ScienceChina, som täcker 2.000 STM-tidskrifter. Två andra som också är stora inom databaser är Vip Information, en statsstödd organisation vars databaser täcker 4.000 tidskrifter, och Tsinghua Tongfang Optical Disc Co. Ltd, en avknoppning från Tsinghua University, vars databaser täcker 7.000 tidskrifter.

Jag uppfattade att antalet kinesiska databaser är 300 st i snäv bemärkelse (dvs av "traditionell" typ, inom t ex STM) (Anm: antalet 300 databaser finns inte nämnt i någon skriftlig dokumentation varför jag här måste reservera mig för dess riktighet), flertalet (72 %) små, bestående av färre än 100.000 poster. Bara 51 % har information mer än 10 år bakåt i tiden. Innehållsmässigt utgörs 60 % av företags- och produkt-databaser, 15 % av databaser inom STM, tidningsdatabaser 12 %, rättsdatabaser 11 % och finansiella databaser 2 %. och 60 % st.

Antalet Internetanvändare i Kina ökar snabbt. I mitten av 2005 var de 103 miljoner, varav över hälften nyttjade bredband. Yrkesmässiga användare utgjorde 11 %, studenter 33 % och lärare 8 %. De mest använda sökmotorerna är Baidu (Kina) och Google. I Peking var andelarna för de främsta enligt följande: Baidu 52 %, Google 33 %, Sohu (Kina) 4.6 %, Sina (Kina) 4 % och Yahoo 3.3 %.

För den som önskar en grundlig översikt av förlagsindustrin i Kina rekommenderade Qiaoqiao Zhang EPS Focus Report (2005) "Business Publishing in China: Issues, Trends and Tactics" och China Dailys 2004-2005 "Annual Report on China's Publishing Industry". Hennes uppsats i Proceedings innehåller en rad ytterligare referenser.

## ***Wikis, bloggar, RSS***

"Wikis, blogs and RSS" tillägnades ett helt spår under hela andra dagen. Uppenbarligen helt rätt: åtminstone att par av sessionerna drog fullt hus – och i något fall släpptes inte ens alla som ville lyssna in! Det tror jag aldrig hänt förut – och det väckte förstås viss irritation bland delegater som betalat dyra pengar kanske just för detta. För mitt vidkommande ställde detta inte till problem eftersom jag valde andra andra aktiviteter. Förlåt ... men det går ju inte att välja allt! Jag hann dock slinka in på de sista minuterna av dagens avslutande paneldiskussion "Monitoring and mining the blogosphere".



*David Tebutt, IWR, Salim Ismail, PubSub, Bradley Horowitz, Yahoo, Loic Le Meur, Six Apart, Fergus Burns, Nooked*

Förutom att ta ett foto hann jag uppfatta att slutsamtalet rörde trenden mot att bloggar börjar bli allt mer kommersiellt intressanta. Det leder till att en del bloggar börjar få rejält med intäkter från annonser, att andra kan börja ta betalt ... Men bloggar kan också användas mer försåtligt för att göra kommersiell "nytta". Kändisar som skriver om sina medievanor kan t ex generera mer trafik till speciella webbplatser, få fler användare till TV-kanaler, få fler köpare av vissa artisters musik ... Ja, nog kan bloggar föda nya affärsmöjligheter för den hågade!

Slutligen nämndes också något om ”strukturerad bloggning”, vilket tycks vara en standard på gång för bloggposter för enkelt utbyte av sådana mellan bloggar, webbplatser etc och för att underlätta sökning och användning av sådana i andra sammanhang.

Märkligt nog - eller möjligen typiskt? - satte de många wiki/blogg/RSS-inslagen på Online bara ett fåtal avtryck i Proceedings. De utgjordes i princip bara av **Marydee Ojalas** “The new face of online: social networking technologies”, som gav en del basala fakta om vad bloggar och wiki är, och **Laurel A. Clydes** ”Wikis as online information sources: collective intelligence or collective idocy” (som dock inte föredrogs!). I Laurels paper står bl a att läsa om (externa) företagswikis, som så sent som hösten 2004 spåddes bli ”årets trend” – vilket inte alls blev fallet. Tvärtom är sådan användning sällsynt. Företagsintern användning av wikis tycks dock vara ganska utbredd.

### ***Growing Global: Lessons learned by an increasingly international information business***

Sessionen “Impact of globalisation on the information market” verkade lovande. En av föredrags-hållarna var ju **William Hann**, Willco/FreePint, UK, mannen bakom det välkända nyhetsbrevet FreePint. Därtill skulle David Lampert från Wolters Kluwer tala om ”The effects of globalisation on the information market”, trevligt för mig med mitt speciella intresse för onlinemarknaden. Men som alltid på konferenser går man på nitar. David dök inte alls upp (och inget paper finns heller i proceedings). Och William Hann pratade inte om FreePint (vilket varit kul) utan ganska tråkigt över ämnet ”Growing Global: Lessons learned by an increasingly international information business”, mot bakgrund av hans erfarenheter av Willco/FreePint, som bl a marknadsför programvaror för publicering och som säljer allt mer på utlandet. Här i alla fall några brottstycken.

Att marknadsföra sig globalt innebär stora möjligheter – men är också en utmaning. Olika tidszoner kräver dygnet-runtservice, olika valutor ger praktiska problem, anpassning till olika kulturer är jobbigt, att kommunicera på olika språk likaså ... Men det är det värt! William gav några tips för den som avser att satsa: Använd enkel teknik. Använd språk och uttryck som är internationellt gångbara (William tipsade om ”skrivregler” på [www.willco.com/support/guides](http://www.willco.com/support/guides), välj Global communication). Interagera med din publik och ta till dig vad de säger. Kommunicera och återkoppla – det ska märkas att du tar till dig synpunkter! William betonade speciellt detta med att interagera: Publicera nyhetsbrev, sprid info med e-mail och RSS, se till att webbplatsen verkligen lever, skapa ett onlineforum, blogga ...

Apropå forum och blogg berättade William att han alltid kollar inläggen innan de släpps för publicering. William berörde också det här med vikten av att bygga sitt varumärke – vilket han verkligen lyckats med beträffande just FreePint. Men valet av just detta namn ångrar han bittert – ty ordet ”free” har inneburit att han inte kunnat använda det för sina abonnemangsbaserade publikationer (VIP och VIP Eye)!

### ***Enterprise search in 2006***

Ett helt spår under den sista konferensdagen ägnades åt ”enterprise search”, det vill säga sökmotorer; det som tidigare kallades IR, information retrieval. Det innebar tre sessioner, varav en paneldiskussion. Noteras kan att bara ett enda av dagens alla föredrag (nämligen mitt eget föredrag, sist) i detta spår återfinns i Proceedings.

I centrum stod **Stephen Arnold** - logiskt nog, mot bakgrund av hans roll som guru inom området sökmotorer. Och det inte bara ur synpunkten ”system” utan också vad gäller webbsökmotorer. Han har ju på senare tid gjort sig bemärkt för sina omfattande studier av Google, vilket bl a resulterat i rapporten ”The Google Legacy” (Infonortics 2005), och ett ”detektivarbete” för att spåra och sammanställa



patent som Google på ett eller annat sätt står bakom, totalt 47 st 2001-2004 och 72 st bara under första halvåret 2005, vilka han nu säljer på en CD. ([www.arnoldit.com](http://www.arnoldit.com))

Stephen är inte bara en väldigt bra estradör som alltid har mycket intressant att komma med – han drar sig inte för att också vara kritisk – även mot de mäktigaste företag. Apropos det senare har Stephen berättat att Googles Larry Page och Sergey Brin vägrade att tala med honom efter att de insett att han nog inte bara tänkte skriva snälla saker om Google.

Och det var just i den senare andan han i sin egenskap av keynote-talare inledde dagen. Jag har faktiskt aldrig förut sett honom så bitsk. Han inledde med att visa en bild med några av de begrepp som flyter omkring i vår bransch: ”Försök att ge en definition av något enda av dessa begrepp på 5 minuter!” (underförstått: det mesta som sägs är flum). (Se bilden till höger!)



Han fortsatte med att ta ner oss på jorden vad gäller det som människan hittills lyckats åstadkomma inom informationssökning. ”Vad gäller bildsökning är vi två tredjedelar på väg – vi är alltså inte ens i nolläge ännu”. Vi arbetar fortfarande med texter - som bara utgör 5 % av allt information.” Slutsats: ”This is early days!”.

Fortsättningen var lika kritisk: ”Det finns inget ’enterprise search’. Det finns bara indexering av delar av innehållet.” Stephen berättade att han undersökt ett hundratal företag vad gäller sökmotoranvändning. Det visade sig att inget av dem använde färre än 5 sökmotorer. ”Varför? Jo, därför att de inte fungerar!”. Från sin egen erfarenhet som konsult berättade han om den federala regeringen i USA (Anm: jag antar att Stephen här avsåg first.gov). ”Först använde de Inktomi. Sedan valde de FAST. Och nu nyligen Vivisimo. Ändå: alla används fortfarande; inte har kastats ut!”

Så var det dags för Stephen att kommentera Autonomys förvärv av Verity (vilket skedde i början av november och som radikalt förändrade marknaden för söksystem): ”De enda som vinner på affären är konsulterna! Deras söksystem är radikalt olika. Autonomy - ”automatic” medan ”Verity kräver

Feature	Autonomy	Verity
Approach	Bayesian	Linguistics
Architecture	IDOL server	Distributed servers
Special features	Discovers	Retrieves
Taxonomy	Autogenerated	Users input categories
Database support	Requires creating documents for the system to index	Can read some database files...slowly
Strength	Certain processes are fast	Strong security and work flow
Weaknesses	Hit and miss relevance	Requires much manual configuration

människor”. Stephen uppgav att ”20 % av Veritys intäkter kommer från OEM. Och under de tre sista åren har Verity fått mer intäkter från konsultverksamhet än från sökmotorer”. Han nämnde också den jakt som konkurrenterna FAST och Google omedelbart inledde på Verity-kunder; FAST genom att erbjuda dem en ”safe passage” över till FAST ESP och Google genom att erbjuda en version av Google Search Appliance gratis för de som kastar ut sin befintliga sökmotor.

”Den sökmotorindustri som tidigare varit relativt statisk har nu ’revolutionerats’ på kort tid. Det finns över 300 sökmotorer på marknaden; hur ska du som är ute efter en sökmotor rimligen kunna jämföra

alla dessa!” Efter att han nämnt några av de mest kända fortsatte Stephen med att säga att ”alla leverantörer påstår att just deras produkt klarar att söka allt. Men sanningen är: Det finns ingen som klarar allt eller passar alla!”. Ta bara en sådan sak som matematiska formler – eller kemiska strukturer: det klarar ingen på ett nöjaktigt sätt.

Vad kommer att hända nu då? Stephen förutsade att det under 2006 kommer att ske ett skifte till ”answers”; det är svaret på sökfrågorna som kommer att stå i centrum. På sikt kommer visualisering att bli allt viktigare. CPU-åtgången ökar exponentiellt – vilket är helt i Googles favör. ”Google är inte längre bara ’sökning’. Stephens slutkläm var (här citerar jag hans korta bidrag i Proceedings: ”Google och Google-inspirerad teknik kommer att omdana företagets IT-användning inom de närmaste 9-12 månaderna”.

## ***New perspectives in enterprise search***

*(Obs - missvisande rubrik!)*

Nu var det nog tänkt att de därpå följande föredragshallarna, en från SAP och en från Google, i egen-  
skap av representanter för två helt olika teknikfilosofier skulle skapa en miljö för en intressant dis-  
kussion, i synnerhet tillsammans med Stephen Arnold. Men så blev det inte alls. Här fick vi oss istället  
till livs rena produktpresentationer, ja i princip produktreklam, för deras respektive produkter. **Karsten  
Hohage** från SAP (Tyskland), vars föredrag hade titeln ”Enterprise search strategies”, motsvarade på  
ett utmärkt sätt den bild Stephen tidigare hade gett oss av ”leverantörer som påstår att deras produkter  
klarar allt”. SAP klarar alltså allt, från klassifiering till sökning i alla typer av information, strukturerad  
eller ostrukturerad. Alla sökfaciliteter man kan tänka sig finns tillhands. Och alla applikationer för  
sökning, med SAP eller annan produkt, kan så nås via SAP Enterprise Portal!

Google-presentationen hölls av **David Bercovich**, Google (USA). Det hade titeln ”Unlocking the po-  
tential of enterprise search in the enterprise” och även om det inte var lika mycket en produktpresen-  
tation som en beskrivning av Googles visioner så gav det inte stort mer än Karstens föredrag. Ingen  
har väl kunnat missa Googles vision om ”att organisera all världens information och göra den univer-  
sellt tillgänglig och användbar”, inte heller att Googles filosofi är ”unified search”, dvs allt ska sökas  
likadant – av Google! På nivån Desktop (dokument, webbsidor, e-mail, IM, PDF, presentationer, kal-  
kyldokument), på nivån Corporate Network (säkert innehåll, CM, fildelning, CRM, Intranät, data-  
baser, företagets egna applikationer) såväl som på den högsta nivån ”Världen” (nyheter, webben,  
forskning, bilder, bloggar, produkter, tryckt material). Googles produkter för respektive nivå är  
Google Desktop, Google Search Appliance och Googles webbsökmotor. Det enda (för mig) intressan-  
ta var uppgiften om att Google Search Appliance har fått över 300 kunder i Europa på ett år.

Stephen såg bara trött ut efter dessa dragningar och någon diskussion att tala om blev det inte. Det  
enda intressanta rörde open-sourceverktyg för enterprise-sökning. Enligt Stephen finns det två ”pålit-  
liga” sådana produkter, nämligen [ht://Dig](http://Dig) och Lucene.

## ***Paneldiskussion - Enterprise search question time***

Nästa session var en paneldiskussion benämnd ”Enterprise search question time”. Den leddes av  
**David Green**, EMEA. I panelen satt **Stephen Arnold**, **David Bercovich**, **Paul Sonderegger** (Endeca,  
USA), **Frank Bilotto** (Vivisimo, USA) och **Tracy Lunt** (DuPont, USA).

En av moderatorns första frågor till panelen var om det på sikt verkligen skulle behövas särskilda sök-  
motorer. T ex kommer sökning att ingå i Microsofts kommande Windows Vista. Är inte sökning på  
väg att bli en ”dussinvara”? Tracy, som är ansvarig för sökning inom DuPont med 55.000 användare  
runt om i världen, svarade att Google visserligen har visat att det är enkelt att söka på webben men att  
man för den skull inte ska inbilla sig att det är lika enkelt inom företag. En viktig orsak är att informa-

tionen inom ett företag ligger utspridd på så många olika ställen och i så många system. Att göra all denna information sökbar är inte enkelt. Därtill krävs kategorisering av informationen; ett område där dagens sökmotorer är synnerligen utvecklade: "kategorisering befinner sig fortfarande på dagis-stadiet". Systemen behöver innehålla så mycket mer än sökordssökning!

### *Övertro och "överlöften"*

Stephen instämde helhjärtat i detta och berättade om den i hans mening oerhört stora okunskapen som finns hos företag i allmänhet beträffande sökning. Företagen tror att söksystemen klarar allt och har enormt höga förväntningar om vad som går att åstadkomma. I själva verket är systemen väldigt utvecklade och de klarar bara en liten del av vad som förväntas av dem. Ingen av de övriga i panelen sade emot, inte heller när Stephen gick vidare med att säga att det inte är konstigt att företagen tror det de tror eftersom sökmotorleverantörerna i sin marknadsföring får det att framstå som att deras produkter klarar allt. "Många magiska fraser får man höra!"

David instämde, och berättade att han som Google-representant fått höra mycket från presumptiva kunder om vad dess konkurrenter utlovat. "De 'överlovar' och försöker därtill mer eller mindre 'lura' kunder att köpa väldigt dyra system för bara enkla behov". Varpå Tracy ryckte in och berättade att de tittat på Google innan de slutligen valde Verity. "Visst är Google väldigt bra för webbaserad information. Men för vår/DuPonts del passade det bara för 10 % av behoven. Det fanns många krav som Google inte alls uppfyllde, t ex säkerhet och anpassningsbarhet". Hon tog bl a upp de problem det innebär att dokument som är skyddade ändå är sökbara och syns i sökresultatet.

Tracy kompletterade med att säga att sökning bara är en av många delar i det mycket större komplexet dokumenthantering. Folk har olika sätt att arbeta, att sköta sin dokumenthantering, att sköta kategorisering etc. Få system klarar att också ta hand om sådant; och inom DuPont byggde man ett eget system för detta eftersom man inte fann något på marknaden som passade. En person bland åhörarna påpekade den komplikation som består i att företag ständigt undergår förändringar, genom uppköp, sammanslagningar, organisationsförändringar, nya marknadsstrategier ... På detta svarade Tracy att "vi är precis i början av processen. DuPont kommer att förändras, men det får man ta."

### *Priser*

Så kom diskussionen in på frågan om priser. Med sin breda erfarenhet av upphandlingar ryckte Stephen omedelbart in och berättade att väldigt få leverantörer berättar om sina priser. "Den enda som har fasta priser är dtSearch". I princip så fungerar det så att leverantörerna frågar den presumptive köparen "vilken budget har du?" Vad det kan leda till exemplifierade han med att berätta att en av hans kunder hade fått betala 30 miljoner kr för att indexera 30-40.000 dokument.

Enligt Stephen finns det en "lucka" mellan gratissystem som Lucene och ht://Dig och det fåtal som klarar stora krav. "Du kan få Ultraseek (anm: Veritys enklare system) gratis men ska du klara ett företags reella behov så måste du ställa upp med en lastbil full med pengar. Det räcker inte bara med systemet utan det krävs också massor av IT-personal". Tracy berättade att deras kostnad för att installera Verity innebar dubbelt så stora kostnader som produkten i sig. Ändå hade man underskattat hårdvarukraven och måste nu utvidga från 4 till 8 serverkluster.

### *Relevans och personalisering*

De representerade produkterna Google, Endeca och Vivisimo satsar alla på att utveckla sina funktioner för relevansbedömning, klustring etc. Paul påpekade att "relevans är det som är bra för mig, precis då jag söker". Detta ledde in diskussionen på "personalisering" (alternativt anpassning till företagets verksamhet) och hur den kan utvecklas. Ett medel är att tillvarata den information som finns i loggar-na; som enligt Stephen utgör "jättelager" av värdefull information som inte utnyttjas och som han menade är ett utmärkt forskningsområde för universiteten att ge sig på. Här fick dock Stephen mot-hugg. David berättade att Google har 500 personer som arbetar med relevansfrågorna och självklart

tittar de också på hur användningen är relaterad till företagets verksamhet. För övrigt får Googles annonsörer på webben gratis tillgång till loggdata. Frank berättade att Vivisimo utvecklar sin teknik för klustring på basis av bl a loggar men bekräftade att han inte visste hur företagen verkligen tillvaratog detta för bättre användning.

Som avslutning fick panelens deltagare möjlighet till några slutord. David framhävde att Google tror på enkelhet, i alla avseenden. Det ska vara enkelt att söka – men det ska också vara enkelt att installera och sköta systemen. ”Våra produkter tar några timmar att installera. Jämför detta med Verity och Autonomy, vilka tar dagar att installera!”. Paul påpekade vikten av att utgå från behoven vid det egna företaget, och då i synnerhet de användargrupper man identifierat som centrala, vid valet av produkt. Tracy menade att man ska engagera företagsledningen och få dem att förstå vad det rör sig om. Den största utmaningen när man inför sökning inom företaget är, enligt Frank, att göra det så enkelt att söka att folk inte blir avskräckta efter första försöket.

Att införa sökning inom ett företag är något av det allra svåraste som finns inom IT-området, sade slutligen Stephen. Gör noggranna studier av vad användarna verkligen behöver - och tro för all del inte att det bara finns FAST, Autonomy och Verity att välja på! Att välja fel system kan innebära att miljontals kr i onödan. Det finns inga genvägar till val av ”rätt” system.

### *Implementing enterprise search at DuPont*

Den allra sista sessionen, ”Selecting and implementing your enterprise search solution”, leddes av **Martin White** i egen hög person. Det bestod av två föredrag vilka redogjorde för två praktikfall. Det första, ”Implementing enterprise search at DuPont”, hölls av **Tracy Lunt**, ansvarig för söksystem inom DuPont, USA.



DuPont har sitt ursprung inom kemi och material men tar numera fram produkter för en lång rad branscher, från livsmedel och textilindustri till sjukvård och elektronik. Omsättningen 2004 var över 200 miljarder kr. Tracy förestår Search Solutions Group, som ingår i DuPonts IT-avdelning. Hon ansvarar för söksystem, informationsarkitektur etc inom DuPonts intranät, som spänner över 70 länder och nyttjas av företagets 55.000 anställda.

För drygt tre år sedan utsågs Verity i det projekt som Tracy ledde till den sökmotor som ska användas inom hela företaget. Innan dess hade DuPont nyttjat ett flertal olika sökmotorer. Man har också en rad andra produkter för informationshantering, bl a Documentum och Oracle. Systemen körs i NT-miljö och man har serverparker som hanterar ”terabytes of data”, varav stora delar utgörs av ostrukturerad information.

Inför valet av Verity upplevde man precis det som andra erfarit: leverantörerna utlovade mirakel! När de sedan skulle ta hand om och indexera dokumenten blev det andra tongångar eftersom de var av en så blandad karaktär. Till exempel passade inte allt in i de taxonomier som fanns att tillgå. Tracy menade att man inte ska inbilla sig att allt kan fungera med automatik; det behövs också manuell taggning. En annan erfarenhet var att man inte ska förlita sig på IT-personal vad gäller att välja söksystem. De har helt andra preferenser och ”bryr sig inte om de potentiella användarnas behov”. Hon nämnde även att det säkert hade varit ”tekniskt” enkelt att ta in Google – men att man då hade missat att tillvarata det arbete man gjorde med att samordna, strukturera och klassificera materialet, vilket också var ett mål i projektet.

Som exempel på vad man uppnått genom projektet var en avsevärd rationalisering av ett arkiv omfattande 250.000 dokument (teknisk dokumentation). Arkivet hanteras numera av 4 personer istället för de 25-40 som behövts tidigare – eftersom en hel del av arbetet sker med automatik.



Däremot har man inte hunnit särskilt långt vad gäller att ta hand om DuPonts information. Trots att man har koll på och har som uppgift att göra all information vid företaget görs sökbar så har man hittills bara indexat 1 % av allt! Det är inte bara en väldig blandning av information – det rör sig också om gigantiska mängder ("det skulle behövas en serverpark i Google-klass för att klara allt!"). En intressant erfarenhet var att många användare upplevde det som lite av en chock när den information de hittills haft "dold" i någon mapp någonstans plötsligt blev sökbar och därmed "synliggjord" för alla i hela företaget. I det läget taggades också många dokument som "hemliga". Här gav oss Tracy rådet att i alla fall se till att åtminstone kontaktperson gjordes "synlig".

Tracy avslutade med "Lessons learned":

*Integration:*

- IT/IS must work closely together
- Do not build a search solution in a technology void
- Be creative – many tools and ways to use them

*Planning:*

- Have a goal (How will it be used? By whom?)
- End-to-End process layout (What effect does each decision have on whole system?)

*On Taxonomies:*

- No software can just "make a Taxonomy"
- No computer system is perfect (but neither is a human system)
- What is success? What is good enough?

### ***Troubles and joys: the evaluation, selection and implementation of an enterprise search engine for a large online information service***

Sessionens sista föredrag hölls av undertecknad **Lars Klasén** i min egenskap av ansvarig för Infodatas projekt för utvärdering och val av ny sökmotor för vårt företag. Projektet startade 2002 och ledde till att avtal i september 2004 tecknades med Fast Search & Transfer (FAST) om anskaffning av dess sökmotor FAST ESP. Den har under 2005 tagits i bruk (i Linux-miljö) för Rättsbanken och dess 40 databaser och är nu på gång även för det statliga personadressregistret SPAR. Fler databaser följer, inom onlinetjänsten InfoTorg såväl som på andra håll inom Infodata.

Bakgrunden till projektet var att Infodata ansåg det dags att ersätta den befintliga sökmotorn Find-It (som vi haft i drift och vidareutvecklat i över 30 år!) med en som uppfyllde en rad nya krav, bland annat avancerad närhetssökning och hantering av XML. Vi ville också slippa egen vidareutveckling. Projektet inleddes med en inventering av produkter på den minst sagt föränderliga marknaden för sökmotorer. Från ca 40 "tänkbara" produkter begränsades antalet så småningom till 12. Efter noggrannare studier av dessa utsågs 7 st som reella kandidater, nämligen BRS/Search (Open Text), RetrievalWare (Convera), FAST, NXT3 (NextPage), Boolware (Softbool), TRIP (TietoEnator) och Verity K2E.

Under 2003 genomfördes tester på plats av några av dessa produkter, med allt från installation och implementering till uppdatering, indexering och sökning i stora databaser. Testerna följde ett uppgjort schema och varje moment dokumenterades – allt för att produkterna skulle kunna jämföras. Kravlistan bara inom "sökning" omfattade över 150 funktioner. Därtill skedde prestanda- och svarstidstester, med eller utan samtidig uppdatering och indexering, och med enstaka såväl som ett stort antal samtidiga "virtuella" användare. Även arkitektur, skalbarhet mm granskades. Därtill intervjuades referenskunder.

Parallellt granskades produkternas roll på marknaden för "enterprise search engines"; en minst sagt rörlig marknad. Vi ville vara så säkra som möjligt på att den valda produkten skulle "överleva" och vidareutvecklas. Att marknaden är rörlig har bekräftats ett antal gånger bara efter avtalets tecknande! Enbart relaterat till kandidatprodukterna har ju NextPages NXT-produkter köpts av FAST (som nu är

marknadstväa med 10-15 % av marknaden) och Verity köpts av Autonomy (som därmed är klar marknadsledare med ca 30 % av marknaden).

På basis av erfarenheterna från projektet gav jag slutligen i mitt föredrag några tips inför val och anskaffning av sökmotor för större företag:

- Kraven på god funktionalitet och god prestanda är delvis motstridiga
- Intervjua även tekniker hos referenskunder
- Ta hänsyn till kostnaderna för hårdvara, mjukvarulicenser och - inte att förgömma - backupmiljöer
- Prisstrukturen skiljer mycket från produkt till produkt - med glöm inte räkna med krav på servrar
- Var beredd på en tidskrävande förankringsprocess hos ledningen
- *Allt* (pris, villkor, ...) är förhandlingsbart – och det gäller alla leverantörer
- Sökmotorer är komplexa och *ingen* är felfri. Problem av ett eller annat slag *kommer* att uppstå
- Konvertering av befintlig information till ny struktur (i vårt fall i XML) kan vara mycket krävande

De viktigaste skälen till valet av FAST ESP var mycket god prestanda, modern arkitektur och goda framtidsutsikter. Men det fanns också nackdelar, bl a att produkten innebar att mycket eget arbete med applikationer mm samt stora indexvolymmer. Men ingen produkt är bäst i allt – och hittills är vi nöjda med vårt val och hoppas kunna köra FAST minst lika länge som vår förra produkt, dvs 30 år ... ☺

## ***Konferensens avslutning***

Den efterföljande frågestunden blev kort eftersom Martin skulle iväg för att säga några avslutande ord i en annan lokal. Vilket han gjorde nästan utan åhörare. Så inte heller i år stod det någon avslutande paneldiskussion på programmet utan det blev faktiskt lite tid för en sista utställningsrunda. Sedan var det slut på Online 2005. Måhända blir det lite högtidligare vid Online Information 2006, vilken äger rum den 28 -30 november 2006. Det är ju den 30:e i ordningen!

/Lars

-----

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